



Étude théorique et numérique des équations différentielles stochastiques rétrogrades

Adrien Richou

► To cite this version:

Adrien Richou. Étude théorique et numérique des équations différentielles stochastiques rétrogrades. Mathématiques [math]. Université Rennes 1, 2010. Français. NNT: . tel-00543719

HAL Id: tel-00543719

<https://theses.hal.science/tel-00543719>

Submitted on 6 Dec 2010

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



THÈSE / UNIVERSITÉ DE RENNES 1
sous le sceau de l'Université Européenne de Bretagne

pour le grade de

DOCTEUR DE L'UNIVERSITÉ DE RENNES 1

Mention : Mathématiques

École doctorale Matisse

présentée par

Adrien Richou

préparée à l'unité de recherche 6625

Institut de recherche mathématique de Rennes
U.F.R. de mathématiques

**Étude théorique
et numérique
des équations différentielles
stochastiques rétrogrades**

**Thèse soutenue à Rennes
le mardi 30 novembre 2010**

devant le jury composé de :

Emmanuel GOBET

Professeur à l'école Polytechnique/rapporteur

Nizar TOUZI

Professeur à l'école Polytechnique/rapporteur

Arnaud DEBUSSCHE

Professeur à l'ENS Cachan - Bretagne/examineur

Denis TALAY

Directeur de recherche à l'INRIA/examineur

Ying HU

Professeur à l'université de Rennes 1/directeur de thèse

Philippe BRIAND

Professeur à l'université de Chambéry/directeur de thèse

To “the Happy Few”

Remerciements

Après avoir passé trois années à faire des mathématiques puis à écrire ce rapport, la rédaction des remerciements s'avère être l'exercice le plus délicat. Il m'est difficile de pouvoir exprimer en quelques lignes toute ma gratitude envers les nombreuses personnes qui m'ont soutenu et qui ont contribué, directement ou indirectement, au travail que vous avez sous les yeux. Je vais tout de même tenter de m'y atteler.

Mes premiers remerciements vont tout naturellement vers mes deux directeurs de thèse Philippe Briand et Ying Hu. Je leur suis extrêmement reconnaissant d'avoir accepté de m'encadrer. Tout au long de ces trois années, ils ont su me prodiguer de bons conseils et m'indiquer les directions à suivre tout en me laissant une très grande autonomie. Leur très grande connaissance de la théorie des EDSRs, leur disponibilité et la confiance qu'ils m'ont accordée font que cette thèse leur doit beaucoup.

Rapporter une thèse est une tâche longue et fastidieuse. Je témoigne donc toute ma gratitude à Emmanuel Gobet et Nizar Touzi qui ont accepté de le faire malgré leur emploi du temps très chargé.

Je suis très honoré que Denis Talay ait accepté de faire partie de mon jury de thèse. En m'accueillant en stage de fin d'étude dans son équipe à Sophia-Antipolis, il a su me donner le goût de la recherche. J'en profite pour remercier également Étienne Tanré qui m'avait encadré pour ce stage et qui m'a fait découvrir le monde des EDSRs. Grâce à eux, je garde un excellent souvenir de mon séjour à l'INRIA.

Merci à Arnaud Debussche pour sa présence dans ce jury et également pour la confiance qu'il m'a accordée en me proposant d'effectuer mon monitorat à l'ENS. Je remercie dans la foulée les autres membres de Ker-Lann : Michel Pierre qui fut un excellent professeur, Grégory Vial dont la gentillesse et les qualités pédagogiques ne sont plus à démontrer et Benoît Cadre avec qui j'ai eu plaisir à travailler.

Merci aux chercheurs de l'IRMAR qui, par leurs qualités humaines et scientifiques, font qu'il est très agréable de travailler dans une telle structure. En particulier, merci à Hélène et Florent pour leur gentillesse et leur bonne humeur permanente, merci à Mihai grâce à qui j'ai eu la chance de faire mon enseignement dans d'excellentes conditions, merci à Nathalie, Lionel, Christophe et Jürgen avec qui j'ai plaisir à discuter, merci à Jean-Baptiste avec qui j'aurais aimé échanger plus. J'ai également une pensée pour le personnel administratif du laboratoire qui est de très grande qualité tant sur le plan professionnel que sur le plan des relations humaines.

Je tiens tout particulièrement à exprimer ma gratitude à Marc Audran, mon professeur de mathématiques spéciales, qui m'a tout simplement appris à réfléchir. En toute honnêteté, je pense que je ne serais pas en train d'écrire ces remerciements aujourd'hui si je n'avais pas suivi ses enseignements.

Les quelques années passées à Rennes m'ont permis de côtoyer de nombreux doctorants qui ont énormément égayé mes journées et mes soirées. Je ne peux commencer ce paragraphe sans évoquer mon « co-bureau » qui a fait que travailler dans le bureau 431 fut un vrai plaisir. Merci Jon pour les nombreuses discussions, les parties de golf, les soirées et ta perpétuelle bonne humeur. Il n'est pas possible de parler du quatrième étage sans également évoquer le bureau 434. Parmi les membres « permanents » je tiens tout particulièrement à remercier Maher, toujours prompt à organiser une soirée improvisée, Nirmal, d'une sagesse incomparable et Alina, quelques grammes de féminité dans ce monde de brute. Il faut rajouter les autres membres du bureau, Adrien, Anna et Julien, ainsi que les habitués des pauses cafés et des mots croisés, Arnaud, champion du monde de blind test, Matthieu, Damien « Flanders », Popoff le « cerbère de la porte », Yiqing, Anjara, Mouton, Julia, et enfin les plus anciens, Thomas, Victor, Ludo, Jimmy, Anne-Claire... N'oublions pas les autres étages avec Gaël, Mathieu, Baptiste, Basile, Jacques, Jean-Louis, avec qui je remercie Mehdi, Yoann, Arnaud, Aurélien, Sébastien, Pierre... J'en oublie sûrement, je propose donc au lecteur éventuellement ignoré de corriger lui même la bourde : je remercie pour

Je profite également de l'occasion pour remercier tous les gens que j'ai pu croiser au cours de ma scolarité et avec qui j'ai encore de très bons contacts. Merci aux Ker-Lannais qui m'ont permis de passer

une très bonne année à l'ENS : Anne, Raphaël, Maud, Mathieu, Élodie, Arnault, Camille... Merci aux Supaeros avec qui je continue de passer régulièrement d'excellents moments : Sébastien, Bastien, Barbe, Marc, Grillon, Nico, Arno, Teuteu, PA... Merci aux Brutions qui ont partagé avec moi de nombreuses années d'internat : Alex, Sico, le Troll, Aymeric, Simon, Mickey, Pierre, Grégoire, Keke... Merci aux plus anciens : Gérald, Mathieu...

Il me semble impensable de terminer ces remerciements sans parler des plus proches. Un très grand merci à ma famille, et ma belle-famille, sur qui j'ai toujours pu compter, en particulier Mamie, Bon-Papa, Bonne-Maman, mes parents qui m'ont toujours fait confiance, mon frère qui sait mieux compter que moi et ma sœur qui sait mieux apprendre à compter que moi. Je ne désespère pas de leur faire tous aimer un jour les maths !

Merci à Barry d'endormir mon fils tous les soirs.

Merci à Jean qui a su me maintenir éveillé la nuit pour terminer ce rapport. Tu m'as permis de découvrir le plus beau des métiers, papa.

Merci Anne-Sophie d'être à mes côtés tous les jours. Ton soutien sans faille et ta joie de vivre permanente sont pour moi le moteur de ma vie.

Table des matières

| | | |
|----------|--|-----------|
| 1 | Introduction | 1 |
| 1.1 | Rappels préliminaires sur la théorie des EDSRs | 1 |
| 1.1.1 | Résultats connus | 2 |
| 1.1.2 | Quelques champs d'application des EDSRs | 3 |
| 1.1.3 | La simulation numérique des EDSRs | 4 |
| 1.2 | Étude des EDSRs ergodiques | 8 |
| 1.2.1 | Résultats connus | 8 |
| 1.2.2 | Résultats nouveaux | 9 |
| 1.2.3 | Quelques résultats complémentaires | 11 |
| 1.3 | Un résultat d'unicité pour les EDSRs quadratiques | 12 |
| 1.3.1 | Résultats connus | 12 |
| 1.3.2 | Résultats nouveaux | 12 |
| 1.4 | Simulation d'EDSRs dont le générateur est à croissance quadratique | 14 |
| 1.4.1 | Résultats connus | 14 |
| 1.4.2 | Résultats théoriques nouveaux | 16 |
| 1.4.3 | Résultats numériques | 18 |
| I | EDSRs ergodiques | 19 |
| 2 | EDSREs et EDPs avec conditions de Neumann au bord | 21 |
| 2.1 | Introduction | 22 |
| 2.2 | EBSDEs with zero Neumann boundary condition | 23 |
| 2.3 | EBSDEs with non-zero Neumann boundary conditions | 27 |
| 2.4 | Study of reflected Kolmogorov processes case | 32 |
| 2.5 | Prob. interpretation of the solution of an elliptic PDE | 34 |
| 2.6 | Optimal ergodic control | 35 |
| 2.7 | EBSDEs on a non-convex bounded set | 38 |
| 2.8 | Appendix | 41 |
| 2.8.1 | Proof of Proposition 2.18. | 41 |
| 2.8.2 | Proof of Corollary 2.19. | 42 |
| 2.8.3 | Proof of Proposition 2.21. | 42 |
| 2.8.4 | Proof of Proposition 2.36. | 43 |
| 3 | Quelques résultats complémentaires sur les EDSREs | 45 |
| 3.1 | Un résultat d'existence et d'unicité pour les EDSRs généralisées | 45 |
| 3.1.1 | Motivation | 45 |
| 3.1.2 | Le résultat d'existence et d'unicité | 45 |
| 3.1.3 | Non application aux EDSREs | 51 |
| 3.2 | Un nouveau théorème d'existence pour les EDSREs | 52 |

| | | |
|------------|---|------------|
| II | Unicité des EDSRs quadratiques | 57 |
| 4 | Unicité des solutions d'EDSRs quadratiques | 59 |
| 4.1 | Introduction | 60 |
| 4.2 | An existence result | 61 |
| 4.3 | A uniqueness result | 62 |
| 4.4 | Application to quadratic PDEs | 67 |
| III | Simulation des EDSRs quadratiques | 75 |
| 5 | Simulation d'EDSRs quadratiques | 77 |
| 5.1 | Introduction | 78 |
| 5.2 | Preliminaries | 79 |
| 5.2.1 | Notations | 79 |
| 5.2.2 | Some results on BMO martingales | 80 |
| 5.2.3 | The backward-forward system | 80 |
| 5.3 | Some useful estimates of Z | 81 |
| 5.3.1 | A first bound for Z | 81 |
| 5.3.2 | A time dependent estimate of Z | 82 |
| 5.3.3 | Zhang's path regularity Theorem | 84 |
| 5.4 | Convergence of a modified time discretization scheme for the BSDE | 86 |
| 5.4.1 | An approximation of the quadratic BSDE | 86 |
| 5.4.2 | Study of the time approximation error $e_3(N, \varepsilon, n)$ | 87 |
| 5.4.3 | Study of the global error $e(N, \varepsilon, n)$ | 92 |
| 5.5 | Some additional results on the time dependent estimate of Z | 96 |
| 5.5.1 | What happens if σ does not depend on time ? | 96 |
| 5.5.2 | Some examples and counterexamples | 97 |
| 5.6 | Appendix | 99 |
| 5.6.1 | Proof of Lemma 5.10. | 99 |
| 5.6.2 | Proof of Lemma 5.11. | 100 |
| 5.6.3 | Proof of Proposition 5.25. | 100 |
| 6 | Résultats numériques | 103 |
| 6.1 | Description de l'algorithme | 104 |
| 6.1.1 | Discretisation temporelle | 104 |
| 6.1.2 | Discretisation spatiale | 105 |
| 6.1.3 | Choix des paramètres | 106 |
| 6.2 | Convergence de l'algorithme | 106 |
| 6.2.1 | Exemples considérés | 106 |
| 6.2.2 | Utilité de la projection pour Z | 108 |
| 6.2.3 | Utilisation de l'algorithme non optimisé | 109 |
| 6.2.4 | Comparaison des différentes grilles de discrétisation | 114 |
| 6.3 | Conclusion | 125 |

Chapitre 1

Introduction

La première partie traite des résultats connus au sujet des EDSRs tandis que les trois parties suivantes sont indépendantes et abordent les résultats nouveaux obtenus au cours de cette thèse. Le début de ce chapitre s'inspire fortement de l'introduction de [19].

1.1 Rappels préliminaires sur la théorie des EDSRs

L'objet de cette thèse est l'étude de certaines équations différentielles stochastiques rétrogrades (notées EDSRs). Commençons donc par rappeler qu'une EDSR est une équation du type

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad 0 \leq t \leq T, \quad (1.1)$$

où $(W_t)_{0 \leq t \leq T}$ est un mouvement brownien d -dimensionnel défini sur un espace de probabilité complet $(\Omega, \mathcal{F}, \mathbb{P})$ dont la filtration naturelle augmentée est notée $(\mathcal{F}_t)_{0 \leq t \leq T}$. Les données d'une telle équation sont d'une part la condition terminale ξ , qui est une variable aléatoire \mathcal{F}_T -mesurable à valeur dans \mathbb{R}^k et d'autre part, le générateur f , fonction aléatoire définie sur $[0, T] \times \Omega \times \mathbb{R}^k \times \mathbb{R}^{k \times d}$, à valeur dans \mathbb{R}^k et mesurable par rapport aux tribus $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}^k) \otimes \mathcal{B}(\mathbb{R}^{k \times d})$ et $\mathcal{B}(\mathbb{R}^k)$, \mathcal{P} désignant la tribu des événements prévisibles. Résoudre une telle équation consiste à trouver un couple de processus $(Y_t, Z_t)_{0 \leq t \leq T}$ adapté par rapport à la filtration engendrée par le mouvement brownien W et vérifiant l'équation (1.1). Donnons une définition plus précise.

Définition 1.1 Une solution de l'EDSR (1.1) est un couple de processus (Y, Z) à valeur dans $\mathbb{R}^k \times \mathbb{R}^{k \times d}$ tel que, Y est continu et adapté, Z est prévisible et \mathbb{P} -p.s., $t \mapsto Z_t$ appartient à $L^2([0, T])$, $t \mapsto f(t, Y_t, Z_t)$ appartient à $L^1([0, T])$. et

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad 0 \leq t \leq T.$$

Avant de poursuivre il peut être bon d'expliquer un peu plus intuitivement la notion d'EDSR. Supposons que l'on souhaite résoudre l'équation différentielle suivante :

$$-\frac{dY_t}{dt} = f(t, Y_t), \quad t \in [0, T],$$

avec la condition terminale $Y_T = \xi$ - d'où la dénomination « rétrograde » - en imposant que, pour tout instant $t \in [0, T]$, Y_t ne dépende pas du futur après t , c'est à dire que le processus Y soit adapté par rapport à la filtration $(\mathcal{F}_t)_{t \in [0, T]}$. Prenons un exemple simple : $f = 0$. On doit donc résoudre l'équation différentielle suivante :

$$-\frac{dY_t}{dt} = 0, \quad t \in [0, T], \quad \text{avec,} \quad Y_T = \xi.$$

Le candidat naturel pour être solution est alors $Y_t = \xi$. Malheureusement ce dernier n'est pas adapté si ξ n'est pas déterministe. Nous pouvons également considérer la meilleure approximation, dans L^2 , adaptée,

donnée par la martingale $Y_t = \mathbb{E}[\xi | \mathcal{F}_t]$. Le théorème de représentation des martingales browniennes nous assure alors l'existence d'un processus Z de carré intégrable et adapté tel que :

$$Y_t = \mathbb{E}[\xi | \mathcal{F}_t] = \mathbb{E}[\xi] + \int_0^t Z_s dW_s.$$

Un simple calcul montre ainsi que

$$Y_t = \xi - \int_t^T Z_s dW_s, \quad \text{i.e.} \quad -dY_t = -Z_t dW_t, \quad \text{avec} \quad Y_T = \xi.$$

La seconde inconnue Z permet donc au processus Y d'être adapté. Pour garder le plus de généralité possible, on peut alors faire dépendre f de Z et ainsi l'équation différentielle initiale devient l'équation différentielle stochastique rétrograde suivante :

$$-dY_t = f(t, Y_t, Z_t)dt - Z_t dW_t, \quad \text{avec,} \quad Y_T = \xi.$$

1.1.1 Résultats connus

Les EDSRs ont été introduites, dans le cas où le générateur est une fonction linéaire, par Bismut [15]. Néanmoins, le point de départ de la théorie des équations rétrogrades est l'article de Pardoux et Peng [74], dans lequel ces deux auteurs considèrent des EDSRs dont le générateur est non linéaire par rapport aux deux variables y et z . Rappelons ce résultat.

Théorème 1.2 *Supposons le générateur f lipschitzien par rapport à (y, z) , uniformément en (t, ω) , et*

$$\mathbb{E} \left[|\xi|^2 + \int_0^T |f(s, 0, 0)|^2 ds \right] < +\infty.$$

Alors l'EDSR (1.1) possède une unique solution (Y, Z) telle que Z soit un processus de carré intégrable.

Il convient de noter que ce résultat peut être démontré en suivant les mêmes idées que pour obtenir l'existence et l'unicité d'une solution pour des équations différentielles ordinaires ou stochastiques : l'article [74] repose sur l'utilisation du lemme de Gronwall mais il est également possible d'utiliser un argument de point fixe (cf [41] par exemple).

Après la publication du théorème cité précédemment, de nombreux auteurs ont cherché à affaiblir les hypothèses sous lesquelles on peut obtenir l'existence et l'unicité de l'EDSR (1.1). La liste complète serait trop longue à énumérer. On peut tout de même diviser les résultats obtenus en deux catégories : ceux valables en toutes dimensions, lorsque Y est à valeur dans \mathbb{R}^k , et ceux qui sont propres à la dimension 1, Y étant alors un processus réel. En ce qui concerne le premier point, les résultats reposent généralement sur l'obtention d'estimations a priori déduites d'une variante du lemme de Gronwall ou bien d'un théorème du point fixe. Signalons notamment l'article [21] qui donne des résultats d'existence et d'unicité pour des EDSRs dont la condition terminale appartient à un L^p pour $p > 1$ et dont le générateur est supposé lipschitzien en z mais seulement « monotone » en y . Cette condition de monotonie s'écrit :

$$\exists \mu, \quad \forall y, y', \quad (y - y') \cdot (f(\cdot, y, \cdot) - f(\cdot, y', \cdot)) \leq \mu |y - y'|^2.$$

Notons également que le cas $p = 1$ est étudié dans ce même article. Pour le second point, à savoir Y à valeurs réelles, les hypothèses peuvent être affaiblies à l'aide du théorème de comparaison¹. En particulier, Lepeltier et San Martín ont construit des solutions maximales et minimales lorsque le générateur est seulement continu en (y, z) et à croissance linéaire. Mais, une des avancées les plus significatives pour cette théorie est due à Kobylanski [60] qui a construit des solutions pour des EDSRs quadratiques, c'est à dire des EDSRs dont le générateur est à croissance quadratique par rapport à la variable z . Pour être plus précis, une EDSR du type (1.1) est dite quadratique lorsque le générateur f satisfait la condition :

$$\forall t, y, z, \quad |f(t, y, z)| \leq \alpha_t + \beta |y| + \frac{\gamma}{2} |z|^2, \quad \mathbb{P}\text{-p.s.},$$

¹Ce dernier permet, si l'on sait comparer deux générateurs f^1, f^2 ainsi que deux conditions terminales ξ^1 et ξ^2 , de comparer les deux solutions associées Y^1 et Y^2 .

où β, γ sont des constantes positives et α est un processus adapté positif satisfaisant la propriété d'intégrabilité suivante :

$$\exists C > 0, \quad \int_0^T \alpha_s ds \leq C, \quad \mathbb{P}\text{-p.s.}$$

La question de l'unicité est également traitée dans ce même article avec des hypothèses supplémentaires. Une partie des résultats de Kobylanski a ensuite été étendue à un cadre plus général par Lepeltier et San Martín [64]. Cette généralisation concerne en fait la croissance du générateur par rapport à la variable y , la croissance par rapport à z étant toujours quadratique. Une des caractéristiques des résultats obtenus dans ces travaux est que la condition terminale ξ est supposée bornée. Or cette condition n'est pas naturelle car si l'on prend l'EDSR de générateur $|z|^2/2$ on peut montrer par une simple transformation exponentielle que $Y_t := \ln \mathbb{E}[e^\xi | \mathcal{F}_t]$ est une solution. On voit ainsi que la bornitude de ξ est une condition a priori trop forte, la simple existence d'un moment exponentiel semblant suffire. Forts de cette remarque, Briand et Hu ont montré dans [24] un résultat d'existence ainsi qu'un résultat d'unicité dans l'article [25]. L'article [36] qui fait l'objet du chapitre 4 consiste justement à améliorer ce résultat d'unicité.

Pour terminer, notons que des résultats supplémentaires peuvent être obtenus pour les EDSRs quadratiques dont la condition terminale est bornée : sous des hypothèses plus fortes concernant la condition terminale et le générateur on peut montrer que l'intégrale stochastique du processus Z par rapport au mouvement brownien W est une martingale à oscillation moyenne bornée, également appelée martingale OMB ou BMO. Ceci permet notamment aux auteurs de [55] d'obtenir un résultat d'unicité. Les auteurs des articles [20] et [1] obtiennent également des estimations sur les solutions qui permettent d'en déduire un résultat de stabilité. Ce type d'argument est essentiel pour obtenir les résultats du chapitre 5.

1.1.2 Quelques champs d'application des EDSRs

Depuis les premiers résultats théoriques obtenus pour les EDSRs, cette théorie n'a cessé de se développer en raison de ses applications dans les domaines des équations aux dérivées partielles (EDPs), de la finance et du contrôle stochastique. Pour ce qui est des applications des EDSRs aux EDPs, il convient d'introduire la classe particulière des EDSRs markoviennes. Pour ces équations, l'aléatoire du générateur et de la condition terminale est donné par une diffusion. Plus précisément (Y, Z) est solution de l'EDSR

$$Y_t = g(X_T) + \int_t^T f(s, X_s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad (1.2)$$

où X est solution de l'EDS

$$X_t = x + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s,$$

et f et g sont des fonctions déterministes. Considérons l'EDP

$$\partial_t u + \mathcal{L}u + f(\cdot, \cdot, u, {}^t \nabla u \sigma) = 0, \quad u(T, \cdot) = g, \quad (1.3)$$

avec \mathcal{L} le générateur du semi-groupe de la diffusion X donné par

$$\mathcal{L}v = \frac{1}{2} \text{Tr}(\sigma^t \sigma \nabla^2 v) + {}^t b \nabla v,$$

et supposons que cette EDP possède une solution suffisamment régulière u . En appliquant la formule d'Itô, on montre finalement que $(u(t, X_t), {}^t \nabla u \sigma(t, X_t))_{0 \leq t \leq T}$ est solution de l'EDSR (1.2). En particulier, la formule $u(0, x) = Y_0$ généralise la formule de Feynman-Kac qui donne une interprétation probabiliste à la solution d'une EDP. Cette relation entre les EDPs et les EDSRs s'utilise dans les deux sens : on peut étudier l'EDSR pour obtenir certaines propriétés de l'EDP associée ou faire le contraire. Une autre application possible de ce type de formule est la simulation numérique de solutions d'EDPs en utilisant des méthodes probabilistes. Peng [73] est le premier à remarquer ce lien entre les EDSRs et les EDPs en travaillant avec des solutions classiques (pour les EDPs). Néanmoins, la théorie des EDSRs permet de travailler de manière naturelle avec des solutions de viscosité (cf [73]). Nous allons rappeler rapidement la définition d'une solution de viscosité. Pour plus de détails, nous renvoyons le lecteur au livre de Barles [10].

Définition 1.3 Une fonction continue u définie sur $[0, T] \times \mathbb{R}^d$ telle que $u(T, \cdot) = h$ est une sous-solution (respectivement sursolution) de viscosité de l'EDP (1.3) si pour toute fonction φ , de classe $C^{1,2}([0, T] \times \mathbb{R}^d)$, on a

$$\partial_t \varphi(t_0, x_0) + \mathcal{L}\varphi(t_0, x_0) - f(t_0, x_0, u(t_0, x_0), {}^t\nabla \varphi(t_0, x_0)) \geq 0, \quad (\text{respectivement } \leq 0)$$

dès que $u - \varphi$ possède un maximum local (respectivement minimum local) au point $(t_0, x_0) \in (0, T) \times \mathbb{R}^d$. Une solution de viscosité est à la fois sous-solution et sursolution de viscosité.

Une formule de Feynman-Kac pour les EDPs semi-linéaires ergodiques avec condition de Neumann est établie dans le chapitre 2 tandis que le cas des EDPs semi-linéaires avec une croissance quadratique par rapport au gradient est étudié dans le chapitre 4.

Un second exemple d'application se trouve dans le domaine du contrôle stochastique. Considérons par exemple le problème de contrôle optimal étudié dans l'article [42] consistant à minimiser la fonction de coût suivante :

$$J(u) = \mathbb{E} \left[\int_0^T h(t, X_t^u, u_t) dt + g(X_T^u) \right],$$

où u désigne un processus de contrôle à valeur dans K un sous ensemble fermé de \mathbb{R}^m , et X^u est le processus défini comme solution de l'EDS suivante :

$$X_t^u = x + \int_0^t b(s, X_s^u) ds + \int_0^t \sigma(s, X_s^u) [dW_s + r(s, X_s^u, u_s) ds]. \quad (1.4)$$

Afin de résoudre ce problème nous introduisons l'hamiltonien du système ψ comme suit :

$$\psi(t, x, z) = \inf_{u \in K} (h(t, x, u) + z \cdot r(t, x, u)).$$

La réponse au problème de contrôle s'exprime à l'aide de la solution (Y, Z) de l'EDSR

$$Y_t = g(X_T) + \int_t^T \psi(s, X_s, Z_s) ds - \int_t^T Z_s dW_s,$$

où X est solution de l'EDS (1.4) avec $r = 0$. On a alors $J \geq Y_0$ et sous certaines conditions on peut montrer l'égalité ainsi qu'exhiber un contrôle optimal² à l'aide de Z . Notons que, sous l'hypothèse que la fonction h est à croissance quadratique par rapport à x et u , on peut montrer que ψ est à croissance quadratique en z et, de ce fait, l'EDSR associée est de type quadratique. D'autres applications des EDSRs au domaine du contrôle optimal sont abordées dans les articles [76] et [40] par exemple. Des problèmes de contrôle ergodique sont étudiés dans le chapitre 2.

Enfin, les mathématiques financières sont également un très grand domaine d'application des EDSRs. N'ayant pas travaillé sur ce type de problématiques au cours de cette thèse, nous renvoyons le lecteur à l'article d'El Karoui, Peng et Quenez [41]. En ce qui concerne le problème plus spécifique des EDSRs quadratiques, on pourra se référer par exemple aux articles [78, 55, 54] ou la thèse [70].

1.1.3 La simulation numérique des EDSRs

Dès lors que des applications pratiques voient le jour, la question de la simulation numérique des EDSRs se pose tout naturellement. Avant de sérier les différentes méthodes proposées dans la littérature pour résoudre ce problème, il est bon de préciser la problématique. Tout d'abord, notons que le problème de la discrétisation et de la simulation d'EDSs est plutôt bien compris et traité³. Par exemple, si le processus X est solution de l'EDS

$$X_t = x + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s, X_s) dW_s,$$

alors la solution la plus simple pour obtenir une approximation en temps discret de X est de considérer le schéma d'Euler

$$\begin{cases} X_{t_n}^n &= x \\ X_{t_{i+1}}^n &= X_{t_i}^n + b(t_i, X_{t_i}^n)(t_{i+1} - t_i) + \sigma(t_i, X_{t_i}^n)(W_{t_{i+1}} - W_{t_i}), \quad i \in \{0, \dots, n-1\}, \end{cases}$$

²Un contrôle optimal est un contrôle qui permet d'atteindre l'infimum de J sur l'ensemble des contrôles admissibles.

³voir par exemple le livre de Kloeden et Platen [59]

avec $0 = t_0 < \dots < t_n = T$ une partition de $[0, T]$. Le processus ainsi obtenu a le bon goût d'être encore adapté. De plus, il est possible d'obtenir assez facilement une vitesse de convergence pour ce schéma. Enfin, un tel algorithme est facilement implantable⁴ sur un ordinateur. Considérons maintenant (Y, Z) la solution de l'EDSR

$$Y_t = \xi + \int_t^T f(s, Y_s, Z_s) ds - \int_t^T Z_s dW_s,$$

que l'on souhaite approcher par un couple (Y^n, Z^n) de processus adaptés à temps discret. Si l'on remplace $\int_{t_i}^{t_{i+1}} f(s, Y_s, Z_s) ds$ par $(t_{i+1} - t_i) f(t_i, Y_{t_i}^n, Z_{t_i}^n)$ alors le théorème de représentation martingale nous assure l'existence d'un processus $(\tilde{Z}_t^n)_{t \in [0, T]}$ tel que l'on ait

$$Y_{t_i}^n = Y_{t_{i+1}}^n + f(t_i, Y_{t_i}^n, Z_{t_i}^n)(t_{i+1} - t_i) - \int_{t_i}^{t_{i+1}} \tilde{Z}_s^n dW_s.$$

Pour $Z_{t_i}^n$ nous allons prendre la meilleure approximation de \tilde{Z}^n par une variable aléatoire \mathcal{F}_{t_i} -mesurable dans $L^2(\Omega \times [t_i, t_{i+1}])$, à savoir

$$Z_{t_i}^n := (t_{i+1} - t_i)^{-1} \mathbb{E} \left[\int_{t_i}^{t_{i+1}} \tilde{Z}_s^n ds \middle| \mathcal{F}_{t_i} \right] = (t_{i+1} - t_i)^{-1} \mathbb{E} \left[Y_{t_{i+1}}^n (W_{t_{i+1}} - W_{t_i}) \middle| \mathcal{F}_{t_i} \right]$$

d'après l'isométrie de l'intégrale d'Itô. Cela nous permet d'obtenir un premier schéma de discrétisation en temps donné par

$$\begin{cases} Y_{t_n}^n = \xi \\ Z_{t_i}^n = (t_{i+1} - t_i)^{-1} \mathbb{E} \left[Y_{t_{i+1}}^n (W_{t_{i+1}} - W_{t_i}) \middle| \mathcal{F}_{t_i} \right] \\ Y_{t_i}^n = \mathbb{E} \left[Y_{t_{i+1}}^n \middle| \mathcal{F}_{t_i} \right] + (t_{i+1} - t_i) f(t_i, Y_{t_i}^n, Z_{t_i}^n), \quad i \in \{0, \dots, n-1\}. \end{cases} \quad (1.5)$$

Un tel schéma est implicite car $Y_{t_i}^n$ est solution d'une équation non linéaire. Néanmoins, si l'on suppose que f est une fonction lipschitz en y alors un argument de point fixe nous assure l'existence et l'unicité d'une solution lorsque h est suffisamment petit car dans ce cas $y \mapsto \mathbb{E} \left[Y_{t_{i+1}}^n \middle| \mathcal{F}_{t_i} \right] + (t_{i+1} - t_i) f(t_i, y, Z_{t_i}^n)$ est une contraction dans $L^2(\mathcal{F}_{t_i})$. Notons qu'il est également possible de modifier légèrement ce schéma pour le rendre explicite en considérant

$$\begin{cases} Y_{t_n}^n = \xi \\ Z_{t_i}^n = (t_{i+1} - t_i)^{-1} \mathbb{E} \left[Y_{t_{i+1}}^n (W_{t_{i+1}} - W_{t_i}) \middle| \mathcal{F}_{t_i} \right] \\ Y_{t_i}^n = \mathbb{E} \left[Y_{t_{i+1}}^n + (t_{i+1} - t_i) f(t_i, Y_{t_{i+1}}^n, Z_{t_i}^n) \middle| \mathcal{F}_{t_i} \right], \quad i \in \{0, \dots, n-1\}. \end{cases} \quad (1.6)$$

Contrairement au schéma d'Euler pour les EDSs, l'algorithme obtenu n'est pas implantable tel quel car il reste à évaluer les espérances conditionnelles.

Discrétisation temporelle et évaluation des espérances conditionnelles. Comme nous allons pouvoir le constater, il existe de nombreux travaux traitant de la simulation des EDSRs. Commençons par traiter ceux qui s'appuient sur la discrétisation temporelle de l'EDSR décrite peu avant. Tout d'abord dans [86] puis dans [84], Zhang établit des propriétés fines de solutions d'EDSRs. En particulier, il obtient un résultat fondamental sur la régularité L^2 de Z . Si on note $0 = t_0 < t_1 < \dots < t_n = T$ une subdivision de $[0, T]$ et on définit

$$\bar{Z}_{t_i} = \frac{1}{t_{i+1} - t_i} \mathbb{E} \left[\int_{t_i}^{t_{i+1}} Z_s ds \middle| \mathcal{F}_{t_i} \right], \quad \forall i \in \{0, \dots, n-1\},$$

alors Zhang montre que sous certaines conditions, on a

$$\sum_{i=0}^{n-1} \mathbb{E} \left[\int_{t_i}^{t_{i+1}} |Z_t - \bar{Z}_{t_i}|^2 dt \right] \leq C \delta_n,$$

⁴L'utilisation du verbe « implanter » en lieu et place du néologisme « implémenter » est un sujet régulièrement débattu même si le second terme a été officiellement adopté par la commission générale de terminologie et de néologie le 20 avril 2007. Pour ma part, je me bornerai à utiliser uniquement le premier terme.

avec δ_n le pas de la subdivision. Ce résultat lui permet d'établir une vitesse de convergence dans [86] pour un schéma moins naturel que (1.5) puis pour le schéma (1.5) dans [84]. Par contre la question de l'évaluation des espérances conditionnelles n'est pas abordée.

Dans [17], Bouchard et Touzi étudient également la vitesse de convergence du schéma implicite (1.5). De plus ils utilisent des techniques du calcul de Malliavin tirées de l'article [16] pour approcher ces espérances conditionnelles. Les auteurs analysent la propagation de l'erreur sur Y mais pas sur Z . En outre, le schéma numérique semble difficile et coûteux à mettre en place en dehors du cas où X est un mouvement brownien ou un mouvement brownien géométrique.

Dans [48], Gobet, Lemor et Warin donnent un schéma totalement implantable et étudient sa vitesse de convergence. Celui-ci se base également sur le schéma implicite (1.5). En se plaçant dans un cadre markovien, la solution $(Y_t, Z_t)_{t \in [0, T]}$ de l'EDSR peut s'exprimer comme un couple de fonctions du temps et de $X : (u(t, X_t), v(t, X_t))_{t \in [0, T]}$. Ainsi, évaluer les espérances conditionnelles revient à évaluer ces fonctions u et v . Pour ce faire, les auteurs proposent d'évaluer leur projection sur un sous-espace vectoriel fonctionnel : typiquement les bases considérées peuvent être une base d'hypercubes de taille fixée ou bien une base de polynômes de degré donné sur ces hypercubes. Ensuite, ces projections sont elles-mêmes évaluées par une méthode de Monte-Carlo : on simule dans un premier temps des réalisations de X , ou d'un processus approchant X , puis on détermine la projection des fonctions u et v à l'aide d'une méthode des moindres carrés. Le schéma étant implicite, il est nécessaire d'ajouter à chaque pas de temps des itérations de Picard afin d'approcher le point fixe. Dans l'article [49], ces trois mêmes auteurs proposent un travail similaire en partant cette fois du schéma explicite (1.6). Notons que ces deux papiers sont tirés de la thèse [63] où de nombreuses simulations numériques sont développées.

Nous avons vu que le schéma proposé par Gobet, Lemor et Warin dans l'article [48] se fait de manière rétrograde en partant de T . Dans l'article [13], Bender et Denk présentent une amélioration de ce schéma en proposant une évaluation des espérances conditionnelles non plus de manière rétrograde mais progressive : l'idée est de considérer les itérations de Picard avant l'évaluation des espérances. Ainsi, si l'on note $(Y^{n,k}, Z^{n,k})$ la k ème itérée de Picard de l'EDSR discrétisée, nous avons

$$\begin{cases} Y^{n,0} = 0 \\ Z^{n,0} = 0 \\ Z_{t_i}^{n,k} = \mathbb{E} \left[\frac{W_{t_{i+1}} - W_{t_i}}{t_{i+1} - t_i} \left(\xi + \sum_{j=i+1}^{n-1} (t_{j+1} - t_j) f(t_j, Y_{t_j}^{n,k-1}, Z_{t_j}^{n,k-1}) \right) \middle| \mathcal{F}_{t_i} \right], & i \in \{0, \dots, n-1\}, \\ Y_{t_i}^{n,k} = \mathbb{E} \left[\xi + \sum_{j=i}^{n-1} (t_{j+1} - t_j) f(t_j, Y_{t_j}^{n,k-1}, Z_{t_j}^{n,k-1}) \middle| \mathcal{F}_{t_i} \right], & i \in \{0, \dots, n\}, \end{cases} \quad (1.7)$$

puis, dans un second temps, les espérances conditionnelles de ce schéma progressif sont évaluées à l'aide des mêmes outils que ceux des articles [48, 49].

Dans l'article [46], Gobet et Labart obtiennent pour des EDSRs markoviennes un développement plus fin de l'erreur d'approximation temporelle du schéma (1.6) sous des conditions fortes de régularité pour f , g , b et σ . Les auteurs trouvent notamment une meilleure vitesse de convergence que celle de [49] lorsque l'on sait simuler de manière exacte le processus progressif X ce qui est le cas par exemple du mouvement brownien, du mouvement brownien géométrique ou bien des processus d'Ornstein-Uhlenbeck.

Tous les articles précédents établissent leurs résultats de vitesse de convergence lorsque g est une fonction lipschitz. Dans l'article [51], Gobet et Makhlof étudient la vitesse de convergence du schéma (1.6) lorsque g est peu régulière. Pour mesurer ce manque de régularité ils introduisent l'espace

$$L_{2,\alpha} := \left\{ g \text{ t.q. } \mathbb{E}[g(X_T)^2] + \sup_{0 \leq t < 1} \frac{\mathbb{E}[g(X_T) - \mathbb{E}[g(X_T)|\mathcal{F}_t]]^2}{(T-t)^\alpha} < +\infty \right\},$$

avec $\alpha \in]0, 1]$. Par exemple, $g \in L_{2,\alpha}$ si cette fonction est α -Hölder, et $g \in L_{2,\frac{1}{2}}$ si c'est l'indicatrice d'un domaine régulier. Les auteurs démontrent que lorsque la grille de discrétisation est uniforme, de pas $h := T/n$, alors la régularité L^2 de Z est de la forme

$$\sum_{i=0}^{n-1} \mathbb{E} \left[\int_{t_i}^{t_{i+1}} |Z_t - \bar{Z}_{t_i}|^2 dt \right] \leq \frac{C}{n^\alpha},$$

et donc la vitesse de convergence du schéma (1.6) est plus faible que dans le cas classique. Ils montrent également qu'il est possible de choisir une grille non uniforme, où les points de discrétisation se concentrent

près de T , afin de retrouver la régularité L^2 de Z classique et donc la vitesse de convergence habituelle du schéma (1.6). La grille non uniforme est donnée par

$$\left\{ t_k := T - T \left(1 - \frac{k}{n} \right)^{\frac{1}{\beta}}, 0 \leq k \leq n \right\}, \quad (1.8)$$

avec $\beta < \alpha$ lorsque $\alpha < 1$.

Marches aléatoires. Une autre idée naturelle pour simuler des EDSRs consiste à remplacer le mouvement brownien W par une marche aléatoire. Pour être précis, considérons une suite $(\varepsilon_n)_{1 \leq k \leq n}$ i.i.d. de variables de Rademacher symétriques, notons $\mathcal{G}_k := \sigma(\varepsilon_1, \dots, \varepsilon_k)$ et posons $h := T/n$ le pas de discrétisation. Le schéma considéré est donc :

$$\begin{cases} Y_{t_n}^n = \xi^n \\ Z_{t_i}^n = h^{-1/2} \mathbb{E} \left[Y_{t_{i+1}}^n \varepsilon_{i+1} \middle| \mathcal{G}_i \right] \\ Y_{t_i}^n = Y_{t_{i+1}}^n + hf(Y_{t_i}^n, Z_{t_i}^n) - \sqrt{h} \varepsilon_{i+1}, \quad i \in \{0, \dots, n-1\}, \end{cases} \quad (1.9)$$

avec ξ^n une variable aléatoire de carré intégrable \mathcal{G} -mesurable qui converge vers ξ dans L^1 . Le schéma est implicite et donc il est bien défini pour n suffisamment grand et f lipschitz en y . De plus, Y_{t_i} est \mathcal{G}_i -mesurable car elle est \mathcal{G}_{i+1} -mesurable et orthogonale à ε_{i+1} . Dans l'article [22], Briand, Delyon et Mémin démontrent la convergence de ce schéma. Ce résultat repose sur un schéma d'approximation de Picard et des résultats concernant la convergence des filtrations tirés de [29]. Malheureusement aucune vitesse de convergence n'est établie. Toldo, dans les articles [82, 83], prolonge ce résultat de convergence aux EDSRs dont le temps terminal est aléatoire.

Dans [66], les auteurs adaptent le schéma (1.9) pour le rendre explicite en s'appuyant sur le schéma d'approximation de Picard. Comme précédemment, aucune vitesse de convergence n'est établie. De plus leur résultat de convergence se restreint au cas où le générateur f ne dépend pas de z .

Dans [27], Cheridito et Stadje prolongent les résultats de l'article [22] au cas des EDSRs dont le générateur a une croissance sous-quadratique par rapport à z . Comme pour la plupart des résultats sur les EDSRs de générateur non lipschitz en z , leur preuve repose sur un théorème de comparaison. Encore une fois, la convergence est prouvée mais aucune vitesse n'est établie.

Systèmes progressifs rétrogrades. De manière plus générale, ces questions concernant la simulation numérique se sont posées pour les systèmes d'équations différentielles stochastiques progressives rétrogrades. Rappelons que la particularité de tels systèmes est d'avoir une diffusion dont les coefficients peuvent dépendre de Y et Z . Plus précisément, la solution d'un système d'EDSPR est un triplé (X, Y, Z) de processus adaptés vérifiant, $\forall t \in [0, T]$,

$$\begin{cases} X_t &= x + \int_0^t b(s, X_s, Y_s, Z_s) ds + \int_0^t \sigma(s, X_s, Y_s) dW_s, \\ Y_t &= g(X_T) + \int_t^T f(s, X_s, Y_s, Z_s) ds - \int_t^T Z_s dW_s. \end{cases} \quad (1.10)$$

Concernant les résultats théoriques sur l'existence et l'unicité de solutions nous renvoyons le lecteur à l'article [33] ainsi qu'à sa bibliographie. Dans [67], Ma, Protter et Yong établissent le lien entre la solution du système (1.10) et la solution, au sens classique du terme, d'une EDP. Les auteurs soumettent alors l'idée de résoudre l'EDSPR en résolvant l'EDP associée. Puis, dans l'article [39], Douglas, Ma et Protter étudient la convergence ainsi que la vitesse d'un schéma aux différences finies pour résoudre l'EDP associée au système (1.10). Malheureusement, ce type de résolution n'est pas envisageable en grande dimension. De plus, les résultats établis nécessitent des conditions de régularité très fortes : notamment g doit être un peu plus de quatre fois différentiable, à dérivées bornées, et σ doit être uniformément elliptique.

Dans l'article [34], Delarue et Menozzi proposent un schéma directement implantable pour résoudre le système (1.10). Leur algorithme s'appuie sur des techniques de quantification⁵ à l'aide de grilles détermi-

⁵Les méthodes de quantification optimale consistent à approcher de manière discrète des vecteurs aléatoires sous la contrainte d'un critère d'optimalité. Par exemple, on peut chercher à remplacer un vecteur aléatoire $X \in L^2(\Omega, \mathbb{R}^d)$ par un vecteur aléatoire \hat{X} prenant au plus N valeurs en minimisant l'erreur L^2 commise :

$$\hat{X} = \arg \min \left\{ \|X - Y\|_2 \mid Y : \Omega \rightarrow \mathbb{R}^d, \text{ mesurable, t.q. Card}(Y(\Omega)) \leq N \right\}.$$

Pour plus de détails, notamment sur les méthodes numériques permettant de résoudre cette problématique, nous renvoyons le lecteur à l'article [72].

nistes : la résolution fait uniquement appel à des méthodes probabilistes contrairement à l'algorithme de l'article [39]. Une vitesse de convergence est établie ce qui permet en pratique de régler « intelligemment » les nombreux paramètres intervenant dans leur schéma. De plus, les hypothèses de régularité sont fortement réduites par rapport à [39].

Autres méthodes. Dans [6], Bally présente un schéma de discrétisation en temps basé sur les temps de sauts d'un processus de Poisson. Cette méthode lui permet de ne pas utiliser la régularité inconnue de Z et de déduire une vitesse de convergence par rapport à l'intensité du processus de Poisson. Cependant, le schéma numérique, pour être implantable, nécessite le calcul d'intégrales de grande dimension ce qui n'est pas traité dans cet article.

Dans [28], Chevance propose un schéma de discrétisation pour des EDSRs markoviennes dont le générateur est indépendant de z , la discrétisation spatiale étant traitée à l'aide de méthodes de quantification optimale. L'algorithme ainsi obtenu est entièrement implantable. Une borne pour l'erreur sur Y est établie mais sous des hypothèses de régularité très fortes sur g et f , tandis que l'erreur sur Z n'est pas étudiée.

Dans [7], Bally et Pagès proposent un schéma numérique basé sur des techniques de quantification pour des EDSRs markoviennes dont le générateur est indépendant de z et qui peuvent être réfléchies. Cette méthode est également développée dans les articles [71, 8]. L'idée est de remplacer X par une chaîne de Markov à temps discret et à nombre d'états finis, ces états étant obtenus par des méthodes de quantification optimale. Une fois estimées les probabilités de transition pour cette chaîne de Markov, les auteurs peuvent facilement implanter une équation de programmation dynamique pour résoudre l'EDSR initiale. Cet algorithme s'avère très efficace lorsque X est un mouvement brownien ou une fonction explicite du mouvement brownien (un mouvement brownien géométrique par exemple). Dans ce cas, les grilles de quantification ainsi que les probabilités de transition sont calculées une fois pour toute. Dans le cas contraire, il peut être nécessaire de les recalculer chaque fois que l'algorithme est relancé ce qui peut être rédhibitoire en terme de temps.

Dans [47], Gobet et Labart proposent un algorithme basé sur le schéma de Picard. Notons (Y^k, Z^k) la k ème itérée de Picard d'une EDSR markovienne définie récursivement par la suite d'EDSRs linéaires

$$\begin{cases} Y^0 = 0, & Z^0 = 0, \\ Y_t^k = g(X_T) + \int_t^T f(s, X_s, Y_s^{k-1}, Z_s^{k-1})ds - \int_t^T Z_s^k dW_s. \end{cases}$$

Sous des conditions de régularité suffisantes, ces EDSRs peuvent être reliées à une suite d'EDPs : en notant $Y_t^k := u_k(t, X_t)$ et $Z_t^k := {}^t\nabla u_k(t, X_t)\sigma(t, X_t)$, nous avons

$$\partial_t u_{k+1} + \mathcal{L}u_{k+1} + f(\cdot, \cdot, u_k, {}^t\nabla u_k\sigma) = 0, \quad u_{k+1}(T, \cdot) = g.$$

Puisque ces EDPs sont linéaires, la formule de Feynman-Kac permet de les résoudre numériquement à l'aide d'une méthode de Monte-Carlo. Pour ce faire les auteurs utilisent une technique de réduction de variance décrite dans [50] ce qui leur permet d'obtenir une bonne vitesse de convergence. De plus, leur schéma d'approximation fournit une solution régulière en espace et en temps. Par contre, afin de pouvoir utiliser la correspondance entre les EDSRs et les EDPs, leur résultat de convergence est établie sous des hypothèses de régularité fortes : σ est uniformément elliptique et g est un peu plus de deux fois différentiable.

1.2 Étude des EDSRs ergodiques

L'objet de la partie I est l'étude d'EDSRs ergodiques, également notées EDSREs. Dans toute cette partie on considère des EDSRs markoviennes à valeurs réelles.

1.2.1 Résultats connus

On suppose que la diffusion X est solution de l'EDS suivante :

$$X_t = x + \int_0^t b(X_s)ds + \int_0^t \sigma(X_s)dW_s.$$

En prenant la condition terminale nulle et en faisant tendre T vers $+\infty$, il est possible de définir une nouvelle classe d'EDSR, appelée EDSR en horizon infini, qui s'écrit

$$Y_t = Y_T + \int_t^T f(X_s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \quad \forall t, T \text{ tels que } 0 \leq t \leq T. \quad (1.11)$$

Un résultat d'existence et d'unicité a été obtenu par Briand et Hu [23] et généralisé par Royer.

Théorème 1.4 *On suppose qu'il existe $K \geq 0$ et $\mu > 0$ tels que*

1. *f est uniformément lipschitz en y et en z , de constante de lipschitz K ,*
2. *f est strictement monotone en y : $\forall x, y, y', z$*

$$(y - y') \cdot (f(x, y, z) - f(x, y', z)) \leq -\mu |y - y'|^2,$$

3. *$|f(\cdot, 0, 0)| \leq K$.*

Alors l'EDSR (1.11) possède une unique solution $(Y, Z)_{t \geq 0}$ progressivement mesurable telle que Y est continu borné et

$$\mathbb{E} \left[\int_0^{+\infty} e^{-2\mu t} |Z_t|^2 dt \right] < +\infty.$$

De plus on a l'estimation $|Y| \leq \frac{K}{\mu}$.

Ces EDSRs markoviennes permettent notamment d'avoir une représentation probabiliste des EDPs elliptiques semi-linéaires du type

$$\mathcal{L}u(x) + f(x, u(x), {}^t \nabla u \sigma(x)) = 0,$$

avec \mathcal{L} le générateur du semi-groupe associé à X . Lorsque f ne dépend plus de y , la condition de stricte monotonie n'est plus vérifiée. Il apparaît alors dans les inconnues une constante, appelée constante d'ergodicité. L'EDSR (1.11) devient :

$$Y_t = Y_T + \int_t^T [f(X_s, Z_s) - \lambda] ds - \int_t^T Z_s dW_s, \quad \forall t, T \text{ tels que } 0 \leq t \leq T. \quad (1.12)$$

La solution de cette EDSR, appelée EDSR ergodique, est alors un triplé (Y, Z, λ) . Ce type d'équation a été introduit dans l'article [43] de Fuhrman, Hu et Tessitore, avec un processus X à valeur dans un espace de Banach, et permet de traiter des problèmes de contrôle ergodique de la forme :

$$I(x, \rho) = \limsup_{T \rightarrow +\infty} \frac{1}{T} \mathbb{E}^{\rho, T} \left[\int_0^T L(X_s^x, \rho_s) ds \right],$$

où ρ est un processus adapté à valeur dans un espace métrique séparable U et $\mathbb{E}^{\rho, T}$ est l'espérance par rapport à la probabilité \mathbb{P}_T^ρ sous laquelle $W_t^\rho = W_t + \int_0^t R(\rho_s) ds$ est un mouvement brownien. Il existe une littérature abondante traitant ces problèmes d'un point de vue analytique en étudiant l'équation de Hamilton-Jacobi-Bellman associée, nous pouvons citer par exemple les papiers d'Arisawa et Lions [4] et d'Arisawa [2], mais le point de vue probabiliste n'a été envisagé que très récemment.

1.2.2 Résultats nouveaux

Le but du chapitre 2 est de construire des EDSRs ergodiques associées à des EDPs ergodiques définies sur un compact, ayant des conditions de Neumann au bord et telles que la constante d'ergodicité apparaît dans la condition au bord. Pour ce faire, nous allons considérer l'EDSR généralisée en horizon infini suivante : pour tous $0 \leq t \leq T < +\infty$,

$$Y_t^x = Y_T^x + \int_t^T [\psi(X_s^x, Z_s^x) - \lambda] ds + \int_t^T [g(X_s^x) - \mu] dK_s^x - \int_t^T Z_s^x dW_s. \quad (1.13)$$

Dans cette équation (X^x, K^x) est la solution d'une EDS réfléchie dans un domaine borné régulier $G = \{\phi > 0\}$ qui est donnée par

$$\begin{cases} X_t^x = x + \int_0^t b(X_s^x) ds + \int_0^t \sigma(X_s^x) dW_s + \int_0^t \nabla \phi(X_s^x) dK_s^x, & t \geq 0; \\ K_t^x = \int_0^t \mathbb{1}_{X_s^x \in \partial G} dK_s^x, & K^x \text{ est croissant.} \end{cases} \quad (1.14)$$

La solution d'une telle équation est un triplé (Y, Z, μ) . μ - appelée « coût ergodique à la frontière » - fait partie des inconnues tandis que λ est simplement un paramètre. Lorsque ψ et g sont des fonctions de y strictement monotones, Pardoux et Zhang [75] ont montré⁶ qu'il existe une unique solution « classique » (Y, Z) . L'apparition de cette nouvelle inconnue s'explique donc par la non monotonie stricte de ces fonctions.

Il existe dans la littérature de nombreux articles qui traitent par des méthodes analytiques les problèmes de contrôles ergodiques avec des conditions frontières. Nous pouvons citer par exemple les travaux de Bensoussan et Frehse [14] dans le cas de conditions homogènes à la frontière, ou bien l'article de Lasry et Lions [62] dans le cas de conditions plus générales. Néanmoins, dans ces publications, la constante μ n'apparaît pas et la constante λ fait partie des inconnues. Il semblerait que les seuls travaux où la constante d'ergodicité μ est présente dans la condition de Neumann sont ceux d'Arisawa [3], de Barles et Da Lio [11] et de Barles et coll. [12]. Il est bon de rappeler une dernière fois que les constantes λ et μ n'ont pas le même rôle : notre résultat principal dit que sous de bonnes hypothèses, pour tout λ il existe une constante μ telle que (1.13) a une solution. À première vue λ ne joue pas de rôle et pourrait être incorporée à ψ , néanmoins notre stratégie de preuve l'utilise : nous montrons tout d'abord que pour tout μ , il existe une unique constante $\lambda := \lambda(\mu)$ pour laquelle (1.13) a une solution et ensuite nous prouvons que $\lambda(\mathbb{R}) = \mathbb{R}$.

Pour être plus précis, nous commençons par traiter le cas d'EDSREs avec une condition de Neumann nulle dans un domaine régulier borné. Comme dans [43], nous introduisons une EDSR en horizon infini de générateur strictement monotone :

$$Y_t^{x,\alpha} = Y_T^{x,\alpha} + \int_t^T [\psi(X_s^x, Z_s^{x,\alpha}) - \alpha Y_s^{x,\alpha}] ds - \int_t^T Z_s^{x,\alpha} dW_s, \quad 0 \leq t \leq T < +\infty, \quad (1.15)$$

avec $\alpha > 0$. Ensuite nous voulons prouver que $(Y^{x,\alpha} - Y_0^{0,\alpha}, Z^{x,\alpha}, \alpha Y_0^{0,\alpha})$ converge, lorsque $\alpha \rightarrow 0$, vers une solution (Y^x, Z^x, λ) de l'EDSRE (1.13) avec $g = 0$ et $\mu = 0$ fixé. Le point clé de la démonstration consiste à montrer que $v_\alpha(x) := Y_0^{x,\alpha}$ est une fonction lipschitz en x uniformément en α . Pour ce faire, on utilise des hypothèses très fortes sur b , σ et le domaine pour que, en moyenne, les trajectoires de X se rejoignent à l'infini avec une vitesse exponentielle : nous supposons que G est convexe et, si l'on note η la constante de dissipativité de b et σ donnée par

$$\eta := \sup_{x,y \in \bar{G}, x \neq y} \left\{ \frac{t(x-y)(b(x)-b(y))}{|x-y|^2} + \frac{\text{Tr}[(\sigma(x)-\sigma(y))^t(\sigma(x)-\sigma(y))]}{2|x-y|^2} \right\},$$

nous supposons également que η est strictement inférieure à une constante négative explicite. Notons qu'une condition de stricte dissipativité est également présente dans l'article [43]. Cette condition étant relativement forte, Debussche, Hu et Tessitore proposent une méthode pour s'en affranchir dans l'article [32].

Dans un deuxième temps nous nous ramenons à une condition de Neumann non nulle en considérant \tilde{v} une fonction vérifiant $\frac{\partial \tilde{v}}{\partial n}(x) + g(x) = \mu, \forall x \in \partial G$. À l'aide du processus $\tilde{v}(X^x)$ nous pouvons alors modifier l'EDSRE initiale (1.13) pour pouvoir obtenir une EDSRE avec condition de Neumann nulle et ainsi appliquer le résultat d'existence précédent. À ce stade de la démonstration nous obtenons les théorèmes 2.9 et 2.10 qui nous assurent que pour tout μ , il existe une unique constante $\lambda := \lambda(\mu)$ pour laquelle (1.13) a une solution.

Enfin, dans un troisième temps, l'étude de la fonction $\mu \mapsto \lambda(\mu)$ montre qu'elle est continue et décroissante. Sous certaines conditions, nous avons également $\lambda(\mu) \xrightarrow{\mu \rightarrow +\infty} -\infty$ et $\lambda(\mu) \xrightarrow{\mu \rightarrow -\infty} +\infty$ ce qui permet de conclure : pour tout λ , il existe une constante μ pour laquelle (1.13) a une solution. Lorsque ψ est bornée (cf théorème 2.13), ce résultat est vrai si nous supposons que

$$\int_{x \in \bar{G}} \mathbb{E}[K_1^x] \nu(dx) > 0$$

avec ν l'unique mesure invariante de X . Cette hypothèse naturelle signifie que la frontière doit être suffisamment vue dans le temps par le processus X . Malheureusement, l'application des EDSREs aux problèmes de contrôle nécessite que ψ soit non borné. Dans ce cas, l'étude est beaucoup plus compliquée : voir les théorèmes 2.15 et 2.20.

⁶En fait ces auteurs ont montré un résultat d'existence et d'unicité lorsque la constante de monotonie μ est supérieure à une valeur explicite. Néanmoins il est possible de généraliser ce résultat pour $\mu > 0$ en appliquant la même démarche que dans l'article [23] : cf. le théorème 3.1.

Une fois obtenus ces résultats d'existence, nous pouvons les appliquer à des problèmes de contrôle optimal dont les coûts sont donnés par

$$I(x, \rho) = \limsup_{T \rightarrow +\infty} \frac{1}{T} \mathbb{E}^{\rho, T} \left[\int_0^T L(X_s^x, \rho_s) ds + \int_0^T [g(X_s^x) - \mu] dK_s^x \right],$$

$$J(x, \rho) = \limsup_{T \rightarrow +\infty} \begin{cases} +\infty & \text{si } \mathbb{E}^{\rho, T}[K_T^x] = 0, \\ \frac{1}{\mathbb{E}^{\rho, T}[K_T^x]} \mathbb{E}^{\rho, T} \left[\int_0^T [L(X_s^x, \rho_s) - \lambda] ds + \int_0^T g(X_s^x) dK_s^x \right] & \text{sinon,} \end{cases}$$

où ρ est un processus adapté à valeur dans un espace métrique séparable U et $\mathbb{E}^{\rho, T}$ est l'espérance par rapport à la probabilité \mathbb{P}_T^ρ sous laquelle $W_t^\rho = W_t + \int_0^t R(\rho_s) ds$ est un mouvement brownien. Avec des hypothèses appropriées et en prenant

$$\psi(x, z) = \inf_{u \in U} \{L(x, u) + zR(u)\}$$

dans (1.13), nous prouvons alors que $\lambda = \inf_\rho I(x, \rho)$ et $\mu = \inf_\rho J(x, \rho)$ où l'infimum est pris sur tous les contrôles admissibles (cf. théorèmes 2.27 et 2.29).

Ce chapitre est organisé ainsi : les EDSREs avec conditions de Neumann nulles sont étudiées dans la partie 2.2, la partie 2.3 est dédiée aux EDSREs avec conditions non nulles, le cas particulier des processus de Kolmogorov réfléchis (pour X) est envisagé dans la partie 2.4, la partie 2.5 donne des liens avec les EDPs ergodiques tandis que la partie 2.6 traite des applications aux problèmes de contrôles ergodiques et, enfin, l'avant-dernière partie fournit quelques résultats supplémentaires lorsque G est non convexe.

1.2.3 Quelques résultats complémentaires

Le but du chapitre 3 est de donner quelques résultats complémentaires au chapitre 2. Ces résultats ne se trouvent pas dans l'article [77]. Tout d'abord, nous étudions la possibilité de montrer de manière plus directe l'existence d'une solution (Y, Z, μ) pour l'EDSRE

$$Y_t^x = Y_T^x + \int_t^T \psi(X_s^x, Z_s^x) ds + \int_t^T [g(X_s^x) - \mu] dK_s^x - \int_t^T Z_s^x dW_s.$$

L'idée consiste à traduire en termes probabilistes la démonstration de l'article [11] ce qui revient à introduire l'EDSR en horizon infini

$$Y_t^{x, \alpha} = Y_T^{x, \alpha} + \int_t^T [\psi(X_s^x, Z_s^{x, \alpha}) - \alpha^2 Y_s^{x, \alpha}] ds + \int_t^T [g(X_s^x) - \alpha Y_s^{x, \alpha}] dK_s^x - \int_t^T Z_s^{x, \alpha} dW_s.$$

En premier lieu, il convient de savoir si cette EDSR possède une solution éventuellement unique. Ce type d'équation a déjà été envisagé, pour Y en dimension quelconque, dans l'article [75], mais l'étude proposée fait apparaître des hypothèses techniques dont nous souhaitons nous affranchir. En utilisant les spécificités de la dimension 1, Briand et Hu [23] proposent un résultat qui va dans ce sens mais se cantonne aux EDSRs non généralisées. Il est néanmoins possible de reprendre leur démonstration pour l'adapter à notre situation : nous obtenons ainsi le théorème 3.1. Puis dans un second temps nous expliquons en quoi les résultats obtenus ne sont pas suffisants pour pouvoir conclure. A posteriori, cette méthode « directe » semble nécessiter des estimations et des résultats de convergence plus fins que ceux employés dans le chapitre 2.

Dans une seconde partie, nous obtenons un nouveau théorème d'existence d'une solution (Y, Z, μ) pour l'EDSRE

$$Y_t^x = Y_T^x + \int_t^T \psi(X_s^x, Z_s^x) ds + \int_t^T [g(X_s^x) - \mu] dK_s^x - \int_t^T Z_s^x dW_s,$$

lorsque ψ est non borné. Dans le chapitre 2, les théorèmes 2.15 et 2.20 proposent déjà ce type de résultat. Néanmoins, le premier nécessite des hypothèses très fortes et le second est démontré lorsque X est un processus de Kolmogorov. L'intérêt de ce nouveau théorème 3.2 est d'avoir des hypothèses proches de celles du théorème 2.20 sans pour autant particulariser le processus X . La preuve repose sur une inégalité de grande déviation tirée de [38].

1.3 Un résultat d'unicité pour les EDSRs quadratiques

La partie II traite de résultats d'unicités pour les EDSRs quadratiques dont la condition terminale est non bornée et le générateur est convexe, ou concave, par rapport à la variable z .

1.3.1 Résultats connus

Dans le chapitre 4 nous nous intéressons à l'EDSR quadratique suivante

$$Y_t = \xi - \int_t^T g(s, Y_s, Z_s) ds + \int_t^T Z_s dW_s, \quad 0 \leq t \leq T. \quad (1.16)$$

Rappelons que le terme EDSR quadratique désigne une EDSR dont le générateur $-g$ a une croissance quadratique par rapport à la variable z . Comme nous l'avons déjà évoqué, les premiers résultats d'existence et d'unicité concernant les EDSRs quadratiques ont été obtenus par Kobylanski [60] puis ont été généralisés par Lepeltier et San Martín [64] en se restreignant au cas où la condition terminale ξ est bornée. Nous avons également vu que cette condition n'est pas naturelle car en prenant le générateur $|z|^2/2$ on peut montrer par une simple transformation exponentielle que $Y_t := \ln \mathbb{E}[e^\xi | \mathcal{F}_t]$ est une solution ce qui laisse supposer que l'existence d'un moment exponentiel pour ξ suffirait. Supposons que g vérifie

$$\forall t, y, z, \quad |g(t, y, z)| \leq \alpha_t + \beta |y| + \frac{\gamma}{2} |z|^2, \quad \mathbb{P}\text{-p.s.},$$

alors Briand et Hu [24, 25] ont montré l'existence d'une solution (Y, Z) sous l'hypothèse

$$\mathbb{E} \left[e^{\gamma e^{\beta T} (|\xi| + \int_0^T \alpha_s ds)} \right] < +\infty.$$

Pour parvenir à ce résultat, ils ont développé une sorte de méthode de localisation pour passer à la limite dans les EDSRs. Bien que cette technique soit assez efficace pour étudier l'existence de solutions, elle n'est d'aucun secours pour étudier l'unicité. Dans l'article [25], ces mêmes auteurs prouvent un résultat d'unicité lorsque g est convexe (ou concave) en la variable z sous l'hypothèse

$$\forall p \in \mathbb{R}^+, \quad \mathbb{E} \left[e^{p(|\xi| + \int_0^T \alpha_s ds)} \right] < +\infty. \quad (1.17)$$

Leur démonstration consiste à montrer un théorème de comparaison en utilisant la méthode de la « différence θ » qui revient, pour deux solutions Y^1 et Y^2 , à estimer $Y^1 - \theta Y^2$ avec $\theta \in (0, 1)$, puis faire tendre θ vers 1. Un tel résultat leur permet de donner une représentation de l'EDP suivante :

$$\partial_t u(t, x) + \mathcal{L}u(t, x) - g(t, x, u(t, x), -^t \nabla u \sigma(t, x)) = 0, \quad u(T, \cdot) = h,$$

lorsque h et g sont sous-quadratiques en x , i.e. :

$$\forall (t, x, y, z) \in [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^{1 \times d}, \quad |h(x)| + |g(t, x, y, z)| \leq f(t, y, z) + C |x|^p,$$

avec $f \geq 0$, $C > 0$ et $p < 2$.

1.3.2 Résultats nouveaux

Le but du chapitre 4 est d'améliorer le résultat d'unicité de [24] en réduisant l'hypothèse (1.17) tout en conservant l'hypothèse de convexité sur g . Vu que notre résultat d'unicité portera sur des EDSRs dont g est lipschitz en y et convexe en z , nous allons supposer que g suit les conditions de croissance suivantes :

$$-\underline{\alpha}_t - r(|y| + |z|) \leq g(t, y, z) \leq \bar{\alpha}_t + \beta |y| + \frac{\gamma}{2} |z|^2.$$

Cette dissymétrie dans la croissance nous permet de réduire encore les hypothèses de moments sur ξ , $\underline{\alpha}$ et $\bar{\alpha}$. Ainsi, le théorème 4.1 nous assure l'existence d'une solution pour l'EDSR (1.16) lorsqu'il existe $p > 1$ tel que

$$\mathbb{E} \left[\exp \left(\gamma e^{\beta T} \xi^- + \gamma \int_0^T \bar{\alpha}_t e^{\beta t} dt \right) + (\xi^+)^p + \left(\int_0^T \underline{\alpha}_t dt \right)^p \right] < +\infty.$$

Ce type de résultat, tirant partie de la croissance dissymétrique du générateur, a déjà été obtenu dans l'article [24] lorsque $\underline{\alpha}$ et $\bar{\alpha}$ sont des constantes.

Passons au résultat principal de ce chapitre, à savoir le résultat d'unicité. Le théorème 4.5 associé au corollaire 4.2 nous assure l'unicité d'une solution pour l'EDSR (1.16) lorsqu'il existe $q > \gamma e^{\beta T}$ et $\varepsilon > 0$ tels que

$$\mathbb{E} \left[\exp \left(q \xi^- + q \int_0^T \bar{\alpha}_t dt \right) + \exp \left(\varepsilon \xi^+ + \varepsilon \int_0^T \underline{\alpha}_t dt \right) \right] < +\infty.$$

Pour montrer ce résultat, nous appliquons une méthode de vérification : nous définissons dans un premier temps un problème de contrôle stochastique et ensuite nous prouvons que Y est la valeur optimale de ce problème de contrôle. Pour être plus précis, nous introduisons la transformée de Legendre-Fenchel de g :

$$f(t, y, q) := \sup_{z \in \mathbb{R}^{1 \times d}} (zq - g(t, y, z)), \quad \forall t \in [0, T], q \in \mathbb{R}^d, y \in \mathbb{R}.$$

Comme g est convexe, elle est également égale à la transformée de Legendre-Fenchel de f :

$$g(t, y, z) := \sup_{q \in \mathbb{R}^d} (zq - f(t, y, q)), \quad \forall t \in [0, T], z \in \mathbb{R}^{1 \times d}, y \in \mathbb{R}.$$

Ainsi nous avons

$$g(s, Y_s, Z_s) \geq Z_s q_s - f(s, Y_s, q_s),$$

ce qui implique

$$Y_t \leq \xi + \int_t^T f(s, Y_s, q_s) ds + \int_t^T Z_s d\tilde{W}_s,$$

avec $d\tilde{W}_s = dW_s - q_s ds$. Imaginons un instant que l'on ait le droit d'appliquer le théorème de Girsanov, alors nous obtenons

$$Y_t \leq \mathbb{E}^{\mathbb{Q}} \left[\xi + \int_t^T f(s, Y_s, q_s) ds \right].$$

Si de plus on a l'existence d'un processus q^* tel que

$$g(s, Y_s, Z_s) = Z_s q_s^* - f(s, Y_s, q_s^*),$$

alors

$$Y_t = \mathbb{E}^{\mathbb{Q}^*} \left[\xi + \int_t^T f(s, Y_s, q_s^*) ds \right].$$

Finalement, le but est de montrer que

$$Y = \text{ess inf}_{q \in \mathcal{A}} Y^q$$

avec

$$Y_t^q = \mathbb{E}^{\mathbb{Q}} \left[\xi + \int_t^T f(s, Y_s, q_s) ds \right].$$

Pour justifier tous ces calculs, il convient de définir correctement un ensemble de contrôles admissibles \mathcal{A} tel que l'on ait le droit d'appliquer le théorème de Girsanov et que l'on ait $q^* \in \mathcal{A}$. Concrètement nous montrons que ces calculs se justifient lorsque T n'est pas trop grand. Dans le cas général, nous nous contentons de coller « bout à bout » des problèmes de contrôle sur des sous-intervalles $[t_i, t_{i+1}]$.

Ce résultat permet également d'améliorer la formule de Feynman-Kac prouvée par Briand et Hu [25] en relâchant un peu les hypothèses sur h et g (cf. théorème 4.10). Ces fonctions peuvent dorénavant avoir une croissance quadratique pas trop grande par rapport à la variable x . Par contre, dans notre cas σ ne dépend plus que du temps.

Ce chapitre est organisé ainsi : un résultat d'existence est prouvé dans la partie 2.2, la partie 2.3 est dédiée au résultat d'unicité, et, enfin, la dernière partie nous permet d'obtenir une formule du type Feynman-Kac.

1.4 Simulation d'EDSRs dont le générateur est à croissance quadratique

La partie III concerne l'étude de la simulation d'EDSRs quadratiques d'un point de vue théorique (chapitre 5) et d'un point de vue numérique (chapitre 6).

1.4.1 Résultats connus

Comme nous l'avons déjà vu dans la partie 1.1.3, la mise au point d'algorithmes efficaces capables de résoudre des EDSRs pour des dimensions raisonnables a été énormément étudié. Néanmoins dans pratiquement tous les travaux cités, le générateur de l'EDSR est une fonction lipschitz en z et cette hypothèse joue un rôle clé dans leurs preuves. Le seul article faisant exception à cette règle est celui de Cheridito et Stadje [27], dans lequel le générateur est supposé être localement lipschitz et avoir une croissance sous-quadratique en la variable z . À l'heure actuelle les résultats théoriques et numériques concernant la simulation d'EDSRs quadratiques sont réduits, nous pouvons donc les passer en revue de manière exhaustive.

Tout d'abord, lorsque le générateur a une forme suffisamment simple, il est possible de contourner le problème en transformant l'EDSR quadratique en une EDSR à croissance linéaire en z . Prenons l'EDSR quadratique la plus simple qui soit,

$$Y_t = \xi + \int_t^T \frac{|Z_s|^2}{2} ds - \int_t^T Z_s dW_s.$$

cette EDSR se simplifie en effectuant une transformation exponentielle, aussi appelée transformation de Cole-Hopf, de cette équation : si nous notons $U = \exp(Y)$ et $V = \exp(Y)Z$ alors (U, V) est solution de l'EDSR

$$U_t = e^\xi - \int_t^T V_s dW_s.$$

Dans ce cas, l'EDSR initiale possède une solution explicite $Y_t = \ln \mathbb{E}[e^\xi | \mathcal{F}_t]$. De manière plus générale, lorsque le générateur est la somme d'un terme quadratique $z \mapsto C|z|^2$ et d'une fonction qui a une croissance linéaire en z , cette méthode ne permet pas forcément de résoudre explicitement l'EDSR mais peut la transformer en une EDSR à croissance linéaire. Ainsi, dans l'article [57], Imkeller, dos Reis et Zhang se placent dans un cadre markovien et supposent que le générateur f est de la forme

$$f(t, x, y, z) = l(t, x, y) + a(t, z) + \frac{\gamma}{2} |z|^2,$$

où l et a sont des fonctions mesurables, l est uniformément lipschitz en x et y , l et a sont continues en t et enfin a est uniformément lipschitz et homogène en z , c'est à dire

$$\forall c \in \mathbb{R}, \forall (s, z) \in [0, T] \times \mathbb{R}^{1 \times d}, \quad a(s, cz) = ca(s, z).$$

Sous cette condition, les auteurs montrent que la transformation exponentielle permet d'obtenir une EDSR lipschitz en x, y et z . Il suffit alors d'appliquer un des algorithmes décrits dans la partie 1.1.3 pour résoudre numériquement la nouvelle EDSR puis appliquer la transformation inverse pour retomber sur la solution de l'EDSR initiale. Dans le même papier [57], les auteurs appliquent cette méthode pour résoudre un problème de maximisation d'utilité exponentielle afin de fixer le prix d'une option sur le cours du pétrole.

Lorsque l'on cherche à résoudre numériquement une EDSR quadratique uniquement dans le but d'obtenir une solution à l'EDP associée, il est également possible de contourner le problème en faisant appel aux EDSRs. Rappelons que résoudre une EDSR revient à déterminer un triplé (X, Y, Z) de processus adaptés vérifiant

$$\begin{cases} X_t = x + \int_0^t b(s, X_s, Y_s, Z_s) ds + \int_0^t \sigma(s, X_s, Y_s) dW_s, \\ Y_t = g(X_T) + \int_t^T f(s, X_s, Y_s, Z_s) ds - \int_t^T Z_s dW_s, \end{cases}$$

l'EDP associée étant

$$\partial_t u + \mathcal{L}u + f(\cdot, \cdot, u, {}^t \nabla u \sigma(\cdot, \cdot, u)) = 0, \quad u(T, \cdot) = g,$$

avec \mathcal{L} donné par

$$\mathcal{L}v = \frac{1}{2} \text{Tr}(\sigma(\cdot, \cdot, v)^t \sigma(\cdot, \cdot, v) \nabla^2 v) + {}^t b(\cdot, \cdot, v, {}^t \nabla v \sigma(\cdot, \cdot, v)) \nabla v.$$

Ainsi, il est parfois possible de transférer la partie quadratique en z du générateur f vers la dérive b sans pour autant modifier l'EDP. Il suffit alors d'utiliser l'une des méthodes de résolution d'EDSPR décrites dans la partie 1.1.3 pour obtenir une solution à l'EDP initiale. Citons par exemple l'article [34] où est résolue numériquement l'équation déterministe KPZ de cette façon. L'inconvénient majeur de cette méthode réside dans le fait qu'elle fait appel aux résultats d'approximations sur les EDSPRs qui, de manière générale, nécessitent des hypothèses plus fortes sur la régularité des coefficients.

L'idée la plus naïve pour résoudre une EDSR quadratique consiste à approcher celle-ci par une EDSR à croissance linéaire que l'on sait résoudre numériquement. Néanmoins, pour pouvoir obtenir une vitesse de convergence, il convient de pouvoir quantifier l'erreur commise lors de l'approximation initiale. Dans le cadre de la théorie des EDSRs quadratiques ce type de résultat n'est pas simple à obtenir et il a fallu attendre l'utilisation de l'outil des martingales OMB pour pouvoir débloquent la situation. Ainsi, récemment, Imkeller et dos Reis proposent dans l'article [56] d'approcher la solution d'une EDSR markovienne (Y, Z) par la solution (Y^N, Z^N) de l'EDSR tronquée suivante :

$$Y_t^N = g(X_T) + \int_t^T f(s, X_s, Y_s^N, h_N(Z_s^N)) ds - \int_t^T Z_s^N dW_s,$$

où $h_N : \mathbb{R}^{1 \times d} \rightarrow \mathbb{R}^{1 \times d}$ est une modification régulière de la projection sur la boule euclidienne de rayon N et de centre 0. Grâce à l'étude de la régularité des trajectoires des processus solutions, les auteurs montrent que pour tout $\beta \geq 1$, l'erreur d'approximation est inférieure à $C_\beta N^{-\beta}$. Comme on suppose que f est localement lipschitzienne par rapport à z , i.e.

$$|f(\cdot, \cdot, \cdot, z) - f(\cdot, \cdot, \cdot, z')| \leq C(1 + |z| + |z'|) |z - z'|,$$

alors l'EDSR tronquée possède un générateur lipschitz par rapport à z . Il est maintenant possible d'avoir une estimation de l'erreur pour l'approximation en temps de l'EDSR approchée à l'aide de n'importe quel schéma de discrétisation temporelle pour des EDSRs de générateur lipschitz. D'après [84, 17, 49], l'erreur quadratique de discrétisation est bornée par C/n avec n le nombre de points de discrétisation. Néanmoins, la constante dépend de la constante de lipschitz du générateur en z et donc de N : elle est de la forme Ce^{CN^2} . Le terme exponentiel résulte de l'utilisation du lemme de Gronwall discret. Ils obtiennent donc l'erreur globale suivante :

$$C_\beta \left(\frac{1}{N^\beta} + \frac{e^{CN^2}}{n} \right), \quad (1.18)$$

Ainsi, lorsque N augmente, n^{-1} doit devenir « petit » très vite et la vitesse de convergence devient mauvaise : si nous prenons $N = \left(\frac{C}{\varepsilon} \log n\right)^{1/2}$ avec $0 < \varepsilon < 1$, alors l'erreur globale devient $C_{\beta, \varepsilon} (\log n)^{-\beta/2}$.

Il est également possible d'approcher une EDSR quadratique par une EDSR à croissance linéaire en approchant g . En effet, lorsque g est une fonction K_g -lipschitzienne, alors, sous certaines conditions, Z est un processus borné par une constante $C(K_g + 1)$ (voir théorème 5.3). L'idée consiste donc à approcher la solution de notre EDSR (Y, Z) par (Y^N, Z^N) , la solution de l'EDSR

$$Y_t^N = g_N(X_T) + \int_t^T f(s, X_s, Y_s^N, Z_s^N) ds - \int_t^T Z_s^N dW_s,$$

où g_N est une approximation K_{g_N} -lipschitz de g . Grâce à l'outil des martingales OMB, nous avons une estimation de l'erreur pour cette approximation : cf. [56, 20] ou proposition 5.9. Par exemple, si g est α -Hölder, nous pouvons obtenir l'erreur d'approximation $CK_{g_N}^{\frac{\alpha}{1-\alpha}}$ (voir proposition 5.17). Finalement, lorsque g est α -Hölder et $K_{g_N} = N$, l'erreur globale s'écrit

$$C \left(\frac{1}{N^{\frac{\alpha}{1-\alpha}}} + \frac{e^{CN^2}}{n} \right), \quad (1.19)$$

et, par conséquent, la vitesse de convergence est également mauvaise : si nous prenons $N = \left(\frac{C}{\varepsilon} \log n\right)^{1/2}$ avec $0 < \varepsilon < 1$, alors l'erreur globale devient $C_\varepsilon (\log n)^{\frac{-\alpha}{2(1-\alpha)}}$.

Concernant la méthode d'approximation d'EDSR consistant à approcher le mouvement brownien par une marche aléatoire, nous avons vu dans la partie 1.1.3 que Cheridito et Stadje, dans l'article [27], ont démontré la convergence de leur schéma lorsque le générateur est à croissance sous-quadratique en z . Ce résultat repose en grande partie sur un théorème de comparaison qu'ils doivent dans un premier temps démontrer. Les auteurs se sont bornés au cas sous-quadratique, sans considérer le cas quadratique, à cause de ce théorème. En effet, ils montrent que pour un exemple simple de schéma, dont le générateur est $|z|^2$, leur théorème de comparaison est faux. Ainsi, il semble vain de vouloir étendre leur résultat au cas quadratique, tout du moins en utilisant les mêmes outils de démonstration.

Pour finir, il est important de noter que les résultats concernant la théorie des EDSRs quadratiques sont très souvent démontrés à l'aide d'un théorème de comparaison, le processus Y de l'EDSR étant obtenu par convergence monotone d'une suite de processus. Ainsi, il semble vain, a priori, de vouloir obtenir des résultats de convergence dans le cadre quadratique pour les schémas reposant sur des itérations de Picard. De même, les schémas implicites semblent moins intéressants à étudier que leurs homologues explicites car en aval se posera le problème de la résolution de l'équation implicite. Néanmoins, il convient de modérer quelque peu nos propos au vu de l'article [81] de Tevzadze dans lequel l'auteur prouve un résultat d'existence et d'unicité pour une certaine classe d'EDSRs quadratiques en utilisant un argument de point fixe. L'idée principale de la démonstration consiste à utiliser pour Z la norme OMB en lieu et place de la norme L^2 habituelle.

1.4.2 Résultats théoriques nouveaux

Le chapitre 5 traite de la discrétisation temporelle d'EDSRs markoviennes dont le générateur a une croissance quadratique par rapport à la variable z . Plus précisément, nous nous intéressons à l'EDSR suivante

$$Y_t = g(X_T) + \int_t^T f(s, X_s, Y_s, Z_s) ds - \int_t^T Z_s dW_s,$$

avec X le processus solution de l'EDS

$$X_t = x + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s) dW_s,$$

g une fonction bornée et f une fonction localement lipschitzienne qui a une croissance quadratique en la variable z , i.e.

$$|f(., ., ., z) - f(., ., ., z')| \leq C(1 + |z| + |z'|) |z - z'|.$$

Les résultats principaux de ce chapitre reposent entièrement sur une majoration déterministe de Z . Nous avons déjà vu que si g est une fonction lipschitzienne alors Z est un processus borné : dans ce cas le générateur de l'EDSR peut être vu comme une fonction lipschitzienne en z et ainsi le problème est résolu. Lorsque l'on suppose g seulement semi-continue inférieurement ou supérieurement alors le théorème 5.4 nous assure, sous certaines hypothèses, la majoration suivante :

$$|Z_t| \leq M_1 + \frac{M_2}{(T-t)^{1/2}}, \quad 0 \leq t < T. \quad (1.20)$$

Ce type d'estimation a tout d'abord été prouvé dans le cas de générateurs à croissance linéaire par Fuhrman et Tessitore dans l'article [44]. Elle est obtenue comme une conséquence de la formule de Bismut-Elworthy ce qui nécessite notamment de supposer l'inversibilité de σ . Dans l'article [35], Delbaen, Hu et Bao obtiennent la même majoration pour des EDSRs dont le générateur a une croissance sur-quadratique par rapport à z . Leur résultat est établi pour des générateurs dépendant uniquement de z , néanmoins leur méthode de démonstration s'adapte très facilement pour pouvoir traiter des générateurs plus généraux : c.f. le théorème 5.4. Notons que l'idée de la démonstration diffère totalement de celle employée par Fuhrman et Tessitore. Ainsi, dans notre cas il n'est pas nécessaire de supposer que σ est inversible, par contre il convient de se restreindre aux fonctions σ déterministes. Notons également que la démonstration fait apparaître une hypothèse technique : on suppose qu'il existe $\lambda \in \mathbb{R}^+$ tel que $\forall \eta \in \mathbb{R}^d$

$$\left| {}^t\eta \sigma(s) [{}^t\sigma(s) {}^t\nabla b(s, x) - {}^t\sigma'(s)] \eta \right| \leq \lambda |{}^t\eta \sigma(s)|^2. \quad (1.21)$$

La partie 5.5 traite plus spécifiquement cette hypothèse. Notamment, lorsque σ ne dépend pas du temps, il est possible de donner des formulations équivalentes à (1.21) plus simples et de montrer que cette hypothèse suffisante n'est pas nécessaire pour obtenir (1.20).

Pour comprendre comment nous pouvons utiliser cette majoration, il convient de rappeler comment on obtient une vitesse de convergence dans le cas classique. Intéressons nous à l'erreur sur le processus Y , l'erreur sur Z s'étudiant de la même façon. La démonstration repose toujours sur l'utilisation du lemme de Gronwall discret : si nous notons $e_i = \mathbb{E} |Y_{t_i} - Y_{t_i}^n|^2$ l'erreur au temps t_i entre la solution Y et la solution de l'EDSR discrétisée en temps Y^n , alors on peut montrer une inégalité de la forme

$$e_i \leq [1 + C(t_{i+1} - t_i) + CK^2(t_{i+1} - t_i)] e_{i+1} + C(t_{i+1} - t_i)^2,$$

avec K la constante de lipschitz de f en z . Revenons maintenant à notre cadre d'étude. Comme nous avons supposé que f est localement lipschitz en z et que nous avons l'estimation (1.20), f est donc une fonction lipschitz en z dont la constante de lipschitz dépend du temps et l'inégalité précédente devient, dans notre cas,

$$e_i \leq \left[1 + C(t_{i+1} - t_i) + C \frac{t_{i+1} - t_i}{T - t_{i+1}} \right] e_{i+1} + C(t_{i+1} - t_i)^2.$$

Notre idée est de choisir une nouvelle grille de discrétisation en temps, non uniforme, telle que $C \frac{t_{i+1} - t_i}{T - t_{i+1}}$ ne dépend pas de i . Cela revient à mettre plus de points de discrétisation près du temps final T que de 0. Pour cela, il suffit de définir les $n + 1$ points de discrétisation par

$$\begin{cases} t_k &= T \left(1 - \left(\frac{\varepsilon}{T} \right)^{k/(n-1)} \right), & 0 \leq k \leq n-1, \\ t_n &= T, \end{cases}$$

avec ε un paramètre. De manière similaire, nous avons vu dans la partie 1.1.3 que Gobet et Makhlof ont utilisé dans l'article [51] des schémas de discrétisation non uniformes pour simuler des EDSRs à croissance linéaire dont la fonction terminale g n'est pas lipschitz : c.f. la grille (1.8). Notons que nous n'obtenons pas la même grille de discrétisation.

À ce stade, plusieurs points techniques sont encore problématiques. Tout d'abord, le générateur f n'est plus une fonction lipschitz en z sur le dernier intervalle de temps $[t_{n-1}, T]$. De plus, nous avons besoin d'adapter à notre cadre d'étude le théorème classique de régularité des trajectoires de Zhang (c.f. [86]). Pour cela, nous commençons par approcher l'EDSR initiale par l'EDSR suivante :

$$Y_t^{N,\varepsilon} = g_N(X_T) + \int_t^T f^\varepsilon(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon}) ds - \int_t^T Z_s^{N,\varepsilon} dW_s, \quad (1.22)$$

avec

$$f^\varepsilon(s, x, y, z) := \mathbb{1}_{s \leq T-\varepsilon} f(s, x, y, z) + \mathbb{1}_{s > T-\varepsilon} f(s, x, y, 0),$$

et g_N une approximation lipschitz de g . Cette approximation de l'EDSR permet à la fois de rendre le générateur globalement lipschitz en z sur l'intervalle $[t_{n-1}, T]$ et de pouvoir appliquer le théorème de régularité 5.7. Enfin, dans certains cas l'intervalle $[t_{n-1}, T]$ est trop grand pour pouvoir obtenir un résultat de convergence, il est alors nécessaire de redécouper cet intervalle de manière uniforme avec au plus n nouveaux points de discrétisation : nous n'en parlerons pas plus ici afin de ne pas compliquer inutilement cette introduction.

Finalement, le schéma de discrétisation temporelle proposé est une simple modification de l'équation de programmation dynamique classique :

$$\begin{cases} Y_{t_n}^{N,\varepsilon,n} &= g_N(X_{t_n}^n) \\ Z_{t_k}^{N,\varepsilon,n} &= \rho_{t_{k+1}} \left(\frac{1}{t_{k+1} - t_k} \mathbb{E} \left[Y_{t_{k+1}}^{N,\varepsilon,n} (W_{t_{k+1}} - W_{t_k}) \middle| \mathcal{F}_{t_k} \right] \right), & 0 \leq k \leq n-1, \\ Y_{t_k}^{N,\varepsilon,n} &= \mathbb{E} \left[Y_{t_{k+1}}^{N,\varepsilon,n} + (t_{k+1} - t_k) f^\varepsilon(t_k, X_{t_k}^n, Y_{t_{k+1}}^{N,\varepsilon,n}, Z_{t_k}^{N,\varepsilon,n}) \middle| \mathcal{F}_{t_k} \right], & 0 \leq k \leq n-1, \end{cases}$$

avec (X^n) le schéma d'Euler associé à X et $\rho_s : \mathbb{R}^{1 \times d} \rightarrow \mathbb{R}^{1 \times d}$ la projection sur la boule

$$B \left(0, M_1 + \frac{M_2}{(T-s)^{1/2}} \right)$$

où M_1 et M_2 sont issues de la majoration (1.20). Grâce à l'estimation (1.20), nous obtenons une vitesse de convergence pour l'approximation temporelle de cette EDSR approchée (voir théorème 5.15). De plus, l'outil des martingales OMB nous donne à nouveau une estimation de l'erreur d'approximation entre (Y, Z) et $(Y^{N,\varepsilon}, Z^{N,\varepsilon})$ (voir proposition 5.9). Finalement, si nous supposons que g est α -Hölder, nous prouvons que l'on peut choisir N et ε pour obtenir une estimation sur l'erreur globale de la forme $Cn^{-\frac{2\alpha}{(2-\alpha)(2+K)-2+2\alpha}}$ (voir théorème 5.20) où K dépend de la constante M_2 définie dans la majoration (1.20) et des constantes liées à f, g, b et σ . Nous obtenons donc une vitesse bien meilleure que (1.19) ou (1.18), mais ce résultat n'est pas forcément très intéressant en pratique puisque la vitesse dépend fortement de K . Cependant, lorsque b est bornée, nous prouvons qu'il est possible de prendre M_2 aussi petit que l'on veut dans l'estimation (1.20). Dans ce cas, nous obtenons une erreur globale inférieure à $C_\eta n^{-(\alpha-\eta)}$, pour tout $\eta > 0$ (voir théorème 5.23).

Ce chapitre est organisé comme suit. Des résultats connus concernant les EDSs et les EDSRs sont rappelés dans la partie 5.2. La partie 5.3 est dédiée aux estimations sur Z : nous donnons une première borne uniforme pour Z , puis une borne dépendant du temps et finalement nous précisons le théorème classique de régularité des trajectoires de Zhang. Puis dans la partie 5.4 nous définissons un schéma de discrétisation temporelle pour des EDSRs quadratiques avec une grille de discrétisation non uniforme et nous obtenons une vitesse de convergence explicite. Enfin, la partie 5.5 étudie plus en détail la majoration (1.20) ainsi que l'hypothèse technique (1.21).

1.4.3 Résultats numériques

Le chapitre 6 traite de la simulation numérique des EDSRs quadratiques. Le but est de confronter à la pratique les résultats théoriques obtenus dans le chapitre 5. Notons que les résultats du chapitre précédent ne concernent que la discrétisation temporelle de l'EDSR. Pour pouvoir avoir un algorithme implantable il convient d'évaluer les espérances conditionnelles. Dans ce but nous utiliserons les méthodes développées dans les articles [48, 49] et la thèse [63] que nous avons déjà décrites dans la partie 1.1.3.

Première partie

EDSRs ergodiques

Chapitre 2

EDSRs ergodiques et EDPs avec conditions de Neumann au bord

Résumé : Nous étudions une nouvelle classe d'équations différentielles stochastiques rétrogrades (EDSRs en abrégé) qui sont reliées à des conditions de Neumann semi-linéaires relatives à des phénomènes ergodiques. La particularité de ces problèmes est que la constante ergodique apparaît dans la condition au bord. Nous étudions l'existence et l'unicité de solutions pour de telles EDSRs ainsi que le lien avec les équations aux dérivées partielles. Nous appliquons ensuite ces résultats à des problèmes de contrôle ergodique optimal.

Mots clés. Équations différentielles stochastiques rétrogrades, contrôle ergodique, conditions de Neumann, équations aux dérivées partielles ergodiques.

Abstract: We study a new class of ergodic backward stochastic differential equations (EBSDEs for short) which is linked with semi-linear Neumann type boundary value problems related to ergodic phenomena. The particularity of these problems is that the ergodic constant appears in Neumann boundary conditions. We study the existence and uniqueness of solutions to EBSDEs and the link with partial differential equations. Then we apply these results to optimal ergodic control problems.

Key words. Backward stochastic differential equations, ergodic control, Neumann boundary conditions, ergodic partial differential equations.

AMS subject classifications. 60H10, 93E20.

This chapter was published in
Stochastic Processes and their Applications
(Volume 119, Issue 9, September 2009, Pages 2945-2969)
under the title:
Ergodic BSDEs and related PDEs with Neumann boundary conditions

2.1 Introduction

In this paper we study the following type of (Markovian) backward stochastic differential equation with infinite horizon that we shall call ergodic BSDEs or EBSDEs for short: for all $0 \leq t \leq T < +\infty$,

$$Y_t^x = Y_T^x + \int_t^T [\psi(X_s^x, Z_s^x) - \lambda] ds + \int_t^T [g(X_s^x) - \mu] dK_s^x - \int_t^T Z_s^x dW_s. \quad (2.1)$$

In this equation $(W_t)_{t \geq 0}$ is a d -dimensional Brownian motion and (X^x, K^x) is the solution to the following forward stochastic differential equation reflected in a smooth bounded domain $G = \{\phi > 0\}$, starting at x and with values in \mathbb{R}^d :

$$\begin{aligned} X_t^x &= x + \int_0^t b(X_s^x) ds + \int_0^t \sigma(X_s^x) dW_s + \int_0^t \nabla \phi(X_s^x) dK_s^x, \quad t \geq 0; \\ K_t^x &= \int_0^t \mathbb{1}_{X_s^x \in \partial G} dK_s^x, \quad K^x \text{ is non decreasing.} \end{aligned} \quad (2.2)$$

Our aim is to find a triple (Y, Z, μ) , where Y, Z are adapted processes taking values in \mathbb{R} and $\mathbb{R}^{1 \times d}$ respectively. $\psi : \mathbb{R}^d \times \mathbb{R}^{1 \times d} \rightarrow \mathbb{R}$ is a given function. Finally, λ and μ are constants: μ , which is called the “boundary ergodic cost”, is part of the unknowns while λ is a given constant.

It is now well known that BSDEs provide an efficient alternative tool to study optimal control problems, see, e.g. [76] or [40]. But to the best of our knowledge, the paper of Fuhrman, Hu and Tessitore [43] is the only one in which BSDE techniques are applied to optimal control problems with ergodic cost functionals that are functionals depending only on the asymptotic behavior of the state (see e.g. costs defined in formulas (2.6) and (2.7) below). This paper deals with the same type of EBSDE as equation (2.1) but without boundary condition (and in infinite dimension): their aim is to find a triple (Y, Z, λ) such that for all $0 \leq t \leq T < +\infty$,

$$Y_t^x = Y_T^x + \int_t^T [\psi(X_s^x, Z_s^x) - \lambda] ds - \int_t^T Z_s^x dW_s, \quad (2.3)$$

where $(W_t)_{t \geq 0}$ is a cylindrical Wiener process in a Hilbert space and X^x is the solution to a forward stochastic differential equation starting at x and with values in a Banach space. In this case, λ is the “ergodic cost”.

There is a fairly large amount of literature dealing by analytical techniques with optimal ergodic control problems without boundary conditions for finite-dimensional stochastic state equations. We just mention papers of Arisawa and Lions [4] and Arisawa [2]. In this framework, the problem is treated through the study of the corresponding Hamilton-Jacobi-Bellman equation. Of course, same questions have been studied in bounded (or unbounded) domains with suitable boundary conditions. For example we refer the reader to Bensoussan and Frehse [14] in the case of homogeneous Neumann boundary conditions and to Lasry and Lions [62] for state-constraint boundary conditions. But in all these works, the constant μ does not appear and the authors are interested in the constant λ instead.

To the best of our knowledge, the only works where the problem of the constant μ appears in the boundary condition of a bounded domain are those of Arisawa [3] and Barles and Da Lio [11]. The purpose of the present paper is to show that backward stochastic differential equations are an alternative tool to treat such “boundary ergodic control problems”. It is worth pointing out that the role of the two constants are different: our main results say that, for any λ and under appropriate hypothesis, there exists a constant μ for which (2.1) has a solution. At first sight λ does not seem to be important and could be incorporated to ψ , but our proof strategy needs it: we first show that, for any μ , there exists a unique constant $\lambda := \lambda(\mu)$ for which (2.1) has a solution and then we prove that $\lambda(\mathbb{R}) = \mathbb{R}$.

To be more precise, we begin to deal with EBSDEs with zero Neumann boundary condition in a bounded convex smooth domain. As in [43], we introduce the class of strictly monotonic backward stochastic differential equations

$$Y_t^{x,\alpha} = Y_T^{x,\alpha} + \int_t^T [\psi(X_s^x, Z_s^{x,\alpha}) - \alpha Y_s^{x,\alpha}] ds - \int_t^T Z_s^{x,\alpha} dW_s, \quad 0 \leq t \leq T < +\infty, \quad (2.4)$$

with $\alpha > 0$ (see [23] or [79]). We then prove that, roughly speaking, $(Y^{x,\alpha} - Y_0^{0,\alpha}, Z^{x,\alpha}, \alpha Y_0^{0,\alpha})$ converge, as $\alpha \rightarrow 0$, to a solution (Y^x, Z^x, λ) of EBSDE (2.3) for all $x \in G$ when (X^x, K^x) is the solution of (2.2) (see Theorem 2.6). When there is non-zero Neumann boundary condition, we consider a function \tilde{v} such that $\frac{\partial \tilde{v}}{\partial n}(x) + g(x) = \mu, \forall x \in \partial G$ and thanks to the process $\tilde{v}(X^x)$ we modify EBSDE (2.1) in order to apply the previous results relating to zero Neumann boundary condition. In Theorems 2.9 and 2.10 we obtain that for any μ , there exists a unique constant $\lambda := \lambda(\mu)$ for which (2.1) has a solution. $\mu \mapsto \lambda(\mu)$ is a continuous decreasing function and, under appropriate hypothesis, we can show that $\lambda(\mu) \xrightarrow{\mu \rightarrow +\infty} -\infty$ and $\lambda(\mu) \xrightarrow{\mu \rightarrow -\infty} +\infty$ which allow us to conclude: see Theorem 2.13 when ψ is bounded and Theorems 2.15 and 2.20 when ψ is bounded in x and Lipschitz in z . All these results are obtained for a bounded convex domain but it is possible to prove some additional results when the domain is not convex.

Moreover we show that we can find a solution of (2.1) such that $Y^x = v(X^x)$ where v is Lipschitz and is a viscosity solution of the elliptic partial differential equation (PDE for short)

$$\begin{cases} \mathcal{L}v(x) + \psi(x, {}^t \nabla v(x) \sigma(x)) = \lambda, & x \in G \\ \frac{\partial v}{\partial n}(x) + g(x) = \mu, & x \in \partial G, \end{cases} \quad (2.5)$$

with

$$\mathcal{L}f(x) = \frac{1}{2} \text{Tr}(\sigma(x) {}^t \sigma(x) \nabla^2 f(x)) + {}^t b(x) \nabla f(x).$$

The above results are then applied to control problems with costs

$$I(x, \rho) = \limsup_{T \rightarrow +\infty} \frac{1}{T} \mathbb{E}^{\rho, T} \left[\int_0^T L(X_s^x, \rho_s) ds + \int_0^T [g(X_s^x) - \mu] dK_s^x \right], \quad (2.6)$$

$$J(x, \rho) = \limsup_{T \rightarrow +\infty} \begin{cases} +\infty & \text{if } \mathbb{E}^{\rho, T}[K_T^x] = 0, \\ \frac{1}{\mathbb{E}^{\rho, T}[K_T^x]} \mathbb{E}^{\rho, T} \left[\int_0^T [L(X_s^x, \rho_s) - \lambda] ds + \int_0^T g(X_s^x) dK_s^x \right] & \text{otherwise,} \end{cases} \quad (2.7)$$

where ρ is an adapted process with values in a separable metric space U and $\mathbb{E}^{\rho, T}$ denotes expectation with respect to \mathbb{P}_T^ρ the probability under which $W_t^\rho = W_t + \int_0^t R(\rho_s) ds$ is a Wiener process on $[0, T]$. $R : U \rightarrow \mathbb{R}^d$ is a bounded function. With appropriate hypothesis and by setting $\psi(x, z) = \inf_{u \in U} \{L(x, u) + zR(u)\}$ in (2.1) we prove that $\lambda = \inf_\rho I(x, \rho)$ and $\mu = \inf_\rho J(x, \rho)$ where the infimum is over all admissible controls.

The paper is organized as follows. In the following section we study EBSDEs with zero Neumann boundary condition. In section 3 we treat the general case of EBSDEs with Neumann boundary condition. In section 4 we study the example of reflected Kolmogorov processes for the forward equation. In section 5 we examine the link between our results on EBSDEs and solutions of elliptic semi-linear PDEs with linear Neumann boundary condition. Section 6 is devoted to optimal ergodic control problems and the last section contains some additional results about EBSDEs on a non-convex bounded set.

2.2 Ergodic BSDEs (EBSDEs) with zero Neumann boundary conditions

Let us first introduce some notations. Throughout this paper, $(W_t)_{t \geq 0}$ will denote a d -dimensional Brownian motion, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For $t \geq 0$, let \mathcal{F}_t denote the σ -algebra $\sigma(W_s; 0 \leq s \leq t)$, augmented with the \mathbb{P} -null sets of \mathcal{F} . The Euclidean norm on \mathbb{R}^d will be denoted by $|\cdot|$. The operator norm induced by $|\cdot|$ on the space of linear operator is also denoted $|\cdot|$. Given a function $f : \mathbb{R}^d \rightarrow \mathbb{R}^k$ we denote $|f|_\infty = \sup_{x \in \mathbb{R}^d} |f(x)|$ and $|f|_{\infty, \mathcal{O}} = \sup_{x \in \mathcal{O}} |f(x)|$ with \mathcal{O} a subset of \mathbb{R}^d .

Let \mathcal{O} be an open connected subset of \mathbb{R}^d . $\mathcal{C}^k(\overline{\mathcal{O}})$, $\mathcal{C}_b^k(\overline{\mathcal{O}})$ and $\mathcal{C}_{lip}^k(\overline{\mathcal{O}})$ will denote respectively the set of real functions of class \mathcal{C}^k on $\overline{\mathcal{O}}$, the set of the functions of class \mathcal{C}^k which are bounded and whose partial derivatives of order less than or equal to k are bounded, and the set of the functions of class \mathcal{C}^k whose partial derivatives of order k are Lipschitz functions.

$\mathcal{M}^2(\mathbb{R}^+, \mathbb{R}^k)$ denotes the space consisting of all progressively measurable processes X , with values in \mathbb{R}^k such that, for all $T > 0$,

$$\mathbb{E} \left[\int_0^T |X_s|^2 ds \right] < +\infty.$$

Throughout this paper we consider EBSDEs where forward equations are stochastic differential equations (SDEs for short) reflected in a bounded subset G of \mathbb{R}^d . To state our results, we use the following assumptions on G :

(G1). There exists a function $\phi \in \mathcal{C}_b^2(\mathbb{R}^d)$ such that $G = \{\phi > 0\}$, $\partial G = \{\phi = 0\}$ and $|\nabla \phi(x)| = 1$, $\forall x \in \partial G$.

(G2). G is a bounded convex set.

If $x \in \partial G$, we recall that $-\nabla \phi(x)$ is the outward unit vector to ∂G in x . We also consider $b : \mathbb{R}^d \mapsto \mathbb{R}^d$ and $\sigma : \mathbb{R}^d \mapsto \mathbb{R}^{d \times d}$, two functions verifying classical assumptions:

(H1). There exist two constants $K_b > 0$ and $K_\sigma > 0$ such that $\forall x, y \in \mathbb{R}^d$,

$$\begin{aligned} |b(x) - b(y)| &\leq K_b |x - y|, \\ \text{and} \\ |\sigma(x) - \sigma(y)| &\leq K_\sigma |x - y|. \end{aligned}$$

We can state the following result, see e.g. [65] Theorem 3.1.

Lemma 2.1 Assume that (G1) and (H1) hold true. Then for every $x \in \overline{G}$ there exists a unique adapted continuous couple of processes $\{(X_t^x, K_t^x); t \geq 0\}$ with values in $\overline{G} \times \mathbb{R}^+$ such that

$$\begin{aligned} X_t^x &= x + \int_0^t b(X_s^x) ds + \int_0^t \sigma(X_s^x) dW_s + \int_0^t \nabla \phi(X_s^x) dK_s^x, \quad t \geq 0; \\ K_t^x &= \int_0^t \mathbb{1}_{X_s^x \in \partial G} dK_s^x, \quad K^x \text{ is non decreasing.} \end{aligned} \quad (2.8)$$

This section is devoted to the following type of BSDE with infinite horizon

$$Y_t^x = Y_T^x + \int_t^T [\psi(X_s^x, Z_s^x) - \lambda] ds - \int_t^T Z_s^x dW_s, \quad 0 \leq t \leq T < +\infty, \quad (2.9)$$

where λ is a real number and is part of the unknowns of the problem and $\psi : \overline{G} \times \mathbb{R}^d \rightarrow \mathbb{R}$ verifies the following general assumptions:

(H2). There exist $K_{\psi,x} \geq 0$ and $K_{\psi,z} \geq 0$ such that

$$|\psi(x, z) - \psi(x', z')| \leq K_{\psi,x} |x - x'| + K_{\psi,z} |z - z'|, \quad \forall x, x' \in \overline{G}, z, z' \in \mathbb{R}^d.$$

We notice that $\psi(\cdot, 0)$ is continuous so there exists a constant M_ψ verifying $|\psi(\cdot, 0)| \leq M_\psi$. As in [43], we start by considering an infinite horizon equation with strictly monotonic drift, namely, for $\alpha > 0$, the equation

$$Y_t^{x,\alpha} = Y_T^{x,\alpha} + \int_t^T [\psi(X_s^x, Z_s^{x,\alpha}) - \alpha Y_s^{x,\alpha}] ds - \int_t^T Z_s^{x,\alpha} dW_s, \quad 0 \leq t \leq T < +\infty. \quad (2.10)$$

Existence and uniqueness have been first studied by Briand and Hu in [23] and then generalized by Royer in [79]. They have established the following result:

Lemma 2.2 Assume that (G1), (H1) and (H2) hold true. Then there exists a unique solution $(Y^{x,\alpha}, Z^{x,\alpha})$ to BSDE (2.10) such that $Y^{x,\alpha}$ is a bounded adapted continuous process and $Z^{x,\alpha} \in \mathcal{M}^2(\mathbb{R}^+, \mathbb{R}^{1 \times d})$. Furthermore, $|Y_t^{x,\alpha}| \leq M_\psi / \alpha$, \mathbb{P} -a.s. for all $t \geq 0$.

We define

$$v_\alpha(x) := Y_0^{x,\alpha}.$$

It is worth noting that $|v_\alpha(x)| \leq M_\psi/\alpha$ and uniqueness of solutions implies that $v_\alpha(X_t^x) = Y_t^{x,\alpha}$. The next step is to show that v_α is uniformly Lipschitz with respect to α . Let

$$\eta := \sup_{x,y \in \overline{G}, x \neq y} \left\{ \frac{t(x-y)(b(x)-b(y))}{|x-y|^2} + \frac{\text{Tr}[(\sigma(x)-\sigma(y))^t(\sigma(x)-\sigma(y))]}{2|x-y|^2} \right\}.$$

We will use the following assumption:

(H3). $\eta + K_{\psi,z}K_\sigma < 0$.

Remark 2.3 When σ is a constant function, (H3) becomes

$$\sup_{x,y \in \overline{G}, x \neq y} \left\{ \frac{t(x-y)(b(x)-b(y))}{|x-y|^2} \right\} < 0,$$

i.e. b is dissipative.

Proposition 2.4 Assume that (G1), (G2), (H1), (H2) and (H3) hold. Then we have, for all $\alpha > 0$ and $x, x' \in \overline{G}$,

$$|v_\alpha(x) - v_\alpha(x')| \leq \frac{K_{\psi,x}}{-\eta - K_{\psi,z}K_\sigma} |x - x'|.$$

Proof. We use a Girsanov argument due to P. Briand and Y. Hu in [23]. Let $x, x' \in \overline{G}$, we set $\tilde{Y}^\alpha := Y^{x,\alpha} - Y^{x',\alpha}$, $\tilde{Z}^\alpha := Z^{x,\alpha} - Z^{x',\alpha}$,

$$\beta_s = \begin{cases} \frac{\psi(X_s^{x'}, Z_s^{x',\alpha}) - \psi(X_s^x, Z_s^{x,\alpha})}{|Z_s^{x',\alpha} - Z_s^{x,\alpha}|^2} (Z_s^{x',\alpha} - Z_s^{x,\alpha}) & \text{if } Z_s^{x',\alpha} - Z_s^{x,\alpha} \neq 0 \\ 0 & \text{otherwise,} \end{cases}$$

$$f_\alpha(s) = \psi(X_s^x, Z_s^{x,\alpha}) - \psi(X_s^{x'}, Z_s^{x,\alpha}),$$

and $\tilde{W}_t = \int_0^t \beta_s ds + W_t$. By hypothesis (H2), β is an \mathbb{R}^d valued adapted process bounded by $K_{\psi,z}$, so we are allowed to apply the Girsanov theorem: for all $T \in \mathbb{R}^+$ there exists a probability \mathbb{Q}_T under which $(\tilde{W}_t)_{t \in [0,T]}$ is a Brownian motion. Then, from equation (2.10) we obtain

$$\tilde{Y}_t^\alpha = \tilde{Y}_T^\alpha - \alpha \int_t^T \tilde{Y}_s^\alpha ds + \int_t^T f_\alpha(s) ds - \int_t^T \tilde{Z}_s^\alpha d\tilde{W}_s, \quad 0 \leq t \leq T. \quad (2.11)$$

Applying Itô's formula to $e^{-\alpha(s-t)} \tilde{Y}_s^\alpha$, we obtain

$$\begin{aligned} \tilde{Y}_t^\alpha &= e^{-\alpha(T-t)} \tilde{Y}_T^\alpha + \int_t^T e^{-\alpha(s-t)} f_\alpha(s) ds - \int_t^T e^{-\alpha(s-t)} \tilde{Z}_s^\alpha d\tilde{W}_s \\ |\tilde{Y}_t^\alpha| &\leq e^{-\alpha(T-t)} \mathbb{E}^{\mathbb{Q}_T} [|\tilde{Y}_T^\alpha| | \mathcal{F}_t] + \int_t^T e^{-\alpha(s-t)} \mathbb{E}^{\mathbb{Q}_T} [|f_\alpha(s)| | \mathcal{F}_t] ds \\ &\leq e^{-\alpha(T-t)} \mathbb{E}^{\mathbb{Q}_T} [|\tilde{Y}_T^\alpha| | \mathcal{F}_t] \\ &\quad + K_{\psi,x} \int_t^T e^{-\alpha(s-t)} \mathbb{E}^{\mathbb{Q}_T} [|X_s^x - X_s^{x'}|^2 | \mathcal{F}_t]^{1/2} ds. \end{aligned}$$

To conclude we are going to use the following lemma whose proof will be given after the proof of Theorem:

Lemma 2.5 Assume that (G1), (G2), (H1), (H2) and (H3) hold. For all $0 \leq t \leq s \leq T$,

$$\mathbb{E}^{\mathbb{Q}^T} \left[|X_s^x - X_s^{x'}|^2 \middle| \mathcal{F}_t \right] \leq e^{2(\eta + K_{\psi, z} K_\sigma)(s-t)} |X_t^x - X_t^{x'}|^2.$$

Furthermore, if σ is constant then, for all $0 \leq t \leq s$, we have

$$|X_s^x - X_s^{x'}| \leq e^{\eta(s-t)} |X_t^x - X_t^{x'}|.$$

From the last inequality, we deduce

$$|\tilde{Y}_t^\alpha| \leq e^{-\alpha(T-t)} \mathbb{E}^{\mathbb{Q}^T} \left[|\tilde{Y}_T^\alpha| \middle| \mathcal{F}_t \right] + K_{\psi, x} |X_t^x - X_t^{x'}| \int_t^T e^{(-\alpha + \eta + K_{\psi, z} K_\sigma)(s-t)} ds,$$

which implies

$$|\tilde{Y}_t^\alpha| \leq e^{-\alpha(T-t)} \frac{M_\psi}{\alpha} + K_{\psi, x} \frac{[1 - e^{(-\alpha + \eta + K_{\psi, z} K_\sigma)(T-t)}]}{\alpha - \eta - K_{\psi, z} K_\sigma} |X_t^x - X_t^{x'}|.$$

Finally, let $T \rightarrow +\infty$ and the claim follows by setting $t = 0$. \square

Proof of Lemma 2.5. Let us apply Itô's formula to $e^{-2(\eta + K_{\psi, z} K_\sigma)(s-t)} |X_s^x - X_s^{x'}|^2$:

$$\begin{aligned} e^{-2(\eta + K_{\psi, z} K_\sigma)(s-t)} |X_s^x - X_s^{x'}|^2 &= |X_t^x - X_t^{x'}|^2 \\ &+ 2 \int_t^s e^{-2(\eta + K_{\psi, z} K_\sigma)(u-t)} \left[(X_u^x - X_u^{x'}) (b(X_u^x) - b(X_u^{x'})) du \right. \\ &+ \frac{1}{2} \text{Tr}[(\sigma(X_u^x) - \sigma(X_u^{x'}))^t (\sigma(X_u^x) - \sigma(X_u^{x'}))] du \\ &+ (X_u^x - X_u^{x'}) \nabla \phi(X_u^x) dK_u^x - (X_u^x - X_u^{x'}) \nabla \phi(X_u^{x'}) dK_u^{x'} \\ &+ (X_u^x - X_u^{x'}) (\sigma(X_u^x) - \sigma(X_u^{x'})) (d\tilde{W}_u - \beta_u du) \\ &\left. - (\eta + K_{\psi, z} K_\sigma) |X_u^x - X_u^{x'}|^2 du \right]. \end{aligned}$$

\overline{G} is a convex set, so ${}^t(x - y) \nabla \phi(x) \leq 0$ for all $(x, y) \in \partial G \times \overline{G}$. Furthermore $|\beta_s| \leq K_{\psi, z}$ and σ is K_σ -Lipschitz. By definition of η we obtain,

$$\begin{aligned} e^{2(-\eta - K_{\psi, z} K_\sigma)(s-t)} |X_s^x - X_s^{x'}|^2 &\leq |X_t^x - X_t^{x'}|^2 \\ &+ 2 \int_t^s e^{-2(\eta + K_{\psi, z} K_\sigma)(s-t)} \left[(X_s^x - X_s^{x'}) (\sigma(X_s^x) - \sigma(X_s^{x'})) \right] d\tilde{W}_s. \end{aligned}$$

Taking the conditional expectation of the inequality we get the first result. To conclude, the stochastic integral is a null function when σ is a constant function. \square

As in [43], we now set

$$\bar{v}_\alpha(x) = v_\alpha(x) - v_\alpha(0),$$

then we have $|\bar{v}_\alpha(x)| \leq \frac{K_{\psi, x}}{-\eta - K_{\psi, z} K_\sigma} |x|$ for all $x \in \overline{G}$ and all $\alpha > 0$, according to Proposition 2.4. Moreover, $\alpha |v_\alpha(0)| \leq M_\psi$ by Lemma 2.2. Thus we can construct by a diagonal procedure a sequence $(\alpha_n)_{n \in \mathbb{N}} \searrow 0$ such that, for all $x \in \overline{G} \cap \mathbb{Q}^d$, $\bar{v}_{\alpha_n}(x) \rightarrow \bar{v}(x)$ and $\alpha_n v_{\alpha_n}(0) \rightarrow \bar{\lambda}$. Furthermore, \bar{v}_α is a $\frac{K_{\psi, x}}{-\eta - K_{\psi, z} K_\sigma}$ -Lipschitz function for every α . So \bar{v} can be extended to a $\frac{K_{\psi, x}}{-\eta - K_{\psi, z} K_\sigma}$ -Lipschitz function defined on the whole \overline{G} , thereby $\bar{v}_{\alpha_n}(x) \rightarrow \bar{v}(x)$ for all $x \in \overline{G}$. Thanks to this construction, we obtain the following theorem which can be proved in the same way as that of Theorem 4.4 in [43].

Theorem 2.6 (Existence of a solution) Assume that (G1), (G2), (H1), (H2) and (H3) hold. Let $\bar{\lambda}$ be the real number and \bar{v} the function constructed previously. We define $\bar{Y}_t^x := \bar{v}(X_t^x)$. Then, there exists a process $\bar{Z}^x \in \mathcal{M}^2(\mathbb{R}^+, \mathbb{R}^{1 \times d})$ such that $\mathbb{P} - a.s.$ $(\bar{Y}^x, \bar{Z}^x, \bar{\lambda})$ is a solution of the EBSDE (2.9) for all $x \in \overline{G}$. Moreover there exists a measurable function $\bar{\zeta} : \mathbb{R}^d \rightarrow \mathbb{R}^{1 \times d}$ such that $\bar{Z}_t^x = \bar{\zeta}(X_t^x)$.

We remark that the solution to EBSDE (2.9) is not unique. Indeed the equation is invariant with respect to addition of a constant to Y . However we have a uniqueness result for λ .

Theorem 2.7 (Uniqueness of λ) Assume that (G1), (H1) and (H2) hold. Let (Y, Z, λ) be a solution of EBSDE (2.9). Then λ is unique among solutions (Y, Z, λ) such that Y is a bounded continuous adapted process and $Z \in \mathcal{M}^2(\mathbb{R}^+, \mathbb{R}^{1 \times d})$.

Proof. We consider (Y, Z, λ) and (Y', Z', λ') as two solutions of the EBSDE (2.9). Let $\tilde{\lambda} = \lambda' - \lambda$, $\tilde{Y} = Y' - Y$ and $\tilde{Z} = Z' - Z$. We have, for all $T \in \mathbb{R}^{+*}$,

$$\tilde{\lambda} = T^{-1} [\tilde{Y}_T - \tilde{Y}_0] + T^{-1} \int_0^T \tilde{Z}_t \beta_t dt - T^{-1} \int_0^T \tilde{Z}_t dW_t$$

with

$$\beta_s = \begin{cases} \frac{\psi(X_s^x, Z_s') - \psi(X_s^x, Z_s)}{|Z_s' - Z_s|^2} (Z_s' - Z_s) & \text{if } Z_s' - Z_s \neq 0 \\ 0 & \text{elsewhere.} \end{cases} \quad (2.12)$$

β is bounded: by the Girsanov theorem there exists a probability measure \mathbb{Q}_T under which $(\tilde{W}_t = W_t - \int_0^t \beta_s ds)_{t \in [0, T]}$ is a Brownian motion. Computing the expectation with respect to \mathbb{Q}_T we obtain

$$|\tilde{\lambda}| = T^{-1} |\mathbb{E}^{\mathbb{Q}_T} [\tilde{Y}_T - \tilde{Y}_0]| \leq \frac{C}{T},$$

because \tilde{Y} is bounded. So we can conclude the proof by letting $T \rightarrow +\infty$. \square

To conclude this section we will show a proposition that will be useful later.

Proposition 2.8 *Assume that (G1) and (H1) hold, G is a bounded set and $\eta < 0$. Then there exists a unique invariant measure ν for the process $(X_t)_{t \geq 0}$.*

Proof. The existence of an invariant measure ν for the process $(X_t)_{t \geq 0}$ is already stated in [80], Theorem 1.21. Let ν and ν' be two invariant measures. For all $f \in C_{lip}^0(\mathbb{R}^d)$ and all $t \in \mathbb{R}^+$ we have

$$\begin{aligned} \left| \int f d\nu - \int f d\nu' \right| &= \left| \int_{\overline{G}} \mathbb{E}[f(X_t^x)] \nu(dx) - \int_{\overline{G}} \mathbb{E}[f(X_t^y)] \nu'(dy) \right| \\ &= \left| \int_{\overline{G}} \int_{\overline{G}} \mathbb{E}[f(X_t^x) - f(X_t^y)] \nu(dx) \nu'(dy) \right| \\ &\leq K_f \int_{\overline{G}} \int_{\overline{G}} \mathbb{E}[|X_t^x - X_t^y|^2]^{1/2} \nu(dx) \nu'(dy), \end{aligned}$$

with K_f the Lipschitz constant of f . We are able to apply Lemma 2.5 with $\psi = 0$: for all $t \in \mathbb{R}^+$

$$\left| \int f d\nu - \int f d\nu' \right| \leq K_f e^{\eta t} \int_{\overline{G}} \int_{\overline{G}} |x - y| \nu(dx) \nu'(dy) \xrightarrow{t \rightarrow +\infty} 0.$$

Then the claim ends by the use of a density argument and the monotone class theorem. \square

2.3 EBSDEs with non-zero Neumann boundary conditions

We are now interested in EBSDEs with non-zero Neumann boundary conditions: we are looking for solutions to the following type of BSDE, for all $0 \leq t \leq T < +\infty$,

$$Y_t^x = Y_T^x + \int_t^T [\psi(X_s^x, Z_s^x) - \lambda] ds + \int_t^T [g(X_s^x) - \mu] dK_s^x - \int_t^T Z_s^x dW_s, \quad (2.13)$$

where λ is a parameter, μ is part of the unknowns of the problem, ψ still verifies (H2) and $g : \overline{G} \rightarrow \mathbb{R}$ verifies the following general assumption:

(F1). $g \in C_{lip}^1(\overline{G})$.

Moreover we use an extra assumption on ϕ :

(G3). $\phi \in C_{lip}^2(\mathbb{R}^d)$.

In this situation we will say that (Y, Z, μ) is a solution of EBSDE (2.13) with λ fixed. But, due to our proof strategy, we will study firstly a modified problem where μ is a parameter and λ is part of the unknowns. In this case, we will say that (Y, Z, λ) is a solution of EBSDE (2.13) with μ fixed. We establish the following result of existence:

Theorem 2.9 (Existence of a solution) *Assume that (G1), (G2), (G3), (H1), (H2), (H3) and (F1) hold true. Then for any $\mu \in \mathbb{R}$ there exist $\lambda \in \mathbb{R}$, $v \in \mathcal{C}_{lip}^0(\bar{G})$, a measurable function $\zeta : \mathbb{R}^d \rightarrow \mathbb{R}^{1 \times d}$ such that, if we define $Y_t^x := v(X_t^x)$ and $Z_t^x := \zeta(X_t^x)$ then $Z^x \in \mathcal{M}^2(\mathbb{R}^+, \mathbb{R}^{1 \times d})$ and $\mathbb{P} - a.s.$ (Y^x, Z^x, λ) is a solution of EBSDE (2.13) with μ fixed, for all $x \in \bar{G}$.*

Proof. Our strategy is to modify EBSDE (2.13) in order to apply Theorem 2.6. According to Theorem 3.2 of [61] there exists $\alpha \in \mathbb{R}$ and $\tilde{v} \in \mathcal{C}_{lip}^2(\bar{G})$ such that

$$\begin{cases} \Delta \tilde{v} - \alpha \tilde{v} = 0 & \text{on } G \\ \frac{\partial \tilde{v}}{\partial n}(x) + g(x) = \mu, & \forall x \in \partial G. \end{cases}$$

We set $\tilde{Y}_t^x = \tilde{v}(X_t^x)$ and $\tilde{Z}_t^x = {}^t \nabla \tilde{v}(X_t^x) \sigma(X_t^x)$. These processes verify for all $0 \leq t \leq T < +\infty$,

$$\tilde{Y}_t^x = \tilde{Y}_T^x - \int_t^T \mathcal{L} \tilde{v}(X_s^x) ds + \int_t^T [g(X_s^x) - \mu] dK_s^x - \int_t^T \tilde{Z}_s^x dW_s,$$

with

$$\mathcal{L} f(x) = \frac{1}{2} \text{Tr}(\sigma(x)^t \sigma(x) \nabla^2 f(x)) + {}^t b(x) \nabla f(x).$$

We now consider the following EBSDE with infinite horizon:

$$\bar{Y}_t^x = \bar{Y}_T^x + \int_t^T [\bar{\psi}(X_s^x, \bar{Z}_s^x) - \lambda] ds - \int_t^T \bar{Z}_s^x dW_s, \quad 0 \leq t \leq T < +\infty, \quad (2.14)$$

with $\bar{\psi}(x, z) = \mathcal{L} \tilde{v}(x) + \psi(x, z + {}^t \nabla \tilde{v}(x) \sigma(x))$. Since σ , ψ and derivatives of \tilde{v} are Lipschitz functions, there exists a constant $K_{\bar{\psi}, x}$ such that we have for all $x, x' \in \bar{G}$ and $z, z' \in \mathbb{R}^d$

$$|\bar{\psi}(x, z) - \bar{\psi}(x', z')| \leq K_{\bar{\psi}, x} |x - x'| + K_{\psi, z} |z - z'|.$$

In particular, $\bar{\psi}$ and ψ are two $K_{\psi, z}$ -Lipschitz functions in z . So we are able to apply Theorem 2.6: there exist $\bar{\lambda} \in \mathbb{R}$, $\bar{v} \in \mathcal{C}_{lip}^0(\bar{G})$ and a measurable function $\bar{\xi} : \mathbb{R}^d \rightarrow \mathbb{R}^{1 \times d}$ such that $(\bar{Y}^x := \bar{v}(X^x), \bar{Z}^x := \bar{\xi}(X^x), \bar{\lambda})$ is a solution of EBSDE (2.14). We set

$$\begin{aligned} Y_t^x &:= \tilde{Y}_t^x + \bar{Y}_t^x = \tilde{v}(X_t^x) + \bar{v}(X_t^x), \\ Z_t^x &:= \tilde{Z}_t^x + \bar{Z}_t^x = {}^t \nabla \tilde{v}(X_t^x) \sigma(X_t^x) + \bar{\xi}(X_t^x). \end{aligned}$$

Then $(Y^x, Z^x, \bar{\lambda})$ is a solution of EBSDE (2.13) linked to μ . □

We have also a result of uniqueness for λ that can be shown exactly as Theorem 2.7:

Theorem 2.10 (Uniqueness of λ) *Assume that (G1), (H1) and (H2) hold. Let (Y, Z, λ) a solution of EBSDE (2.13) with μ fixed. Then λ is unique among solutions (Y, Z, λ) such that Y is a bounded continuous adapted process and $Z \in \mathcal{M}^2(\mathbb{R}^+, \mathbb{R}^{1 \times d})$.*

Thanks to the uniqueness we can define the map $\mu \mapsto \lambda(\mu)$ and study its properties.

Proposition 2.11 *Assume that (G1), (G2), (G3), (H1), (H2), (H3) and (F1) hold true. Then $\lambda(\mu)$ is a decreasing continuous function on \mathbb{R} .*

Proof. Let (Y^x, Z^x, λ) and $(\tilde{Y}^x, \tilde{Z}^x, \tilde{\lambda})$ be two solutions of (2.13) linked to μ and $\tilde{\mu}$. We set $\bar{Y}^x := \tilde{Y}^x - Y^x$ and $\bar{Z}^x := \tilde{Z}^x - Z^x$. These processes verify for all $T \in \mathbb{R}^+$

$$\bar{Y}_0^x = \bar{Y}_T^x + \int_0^T [\psi(X_s^x, \tilde{Z}_s^x) - \psi(X_s^x, Z_s^x)] ds + [\lambda - \tilde{\lambda}]T + [\mu - \tilde{\mu}]K_T^x - \int_0^T \bar{Z}_s^x dW_s. \quad (2.15)$$

As usual, we set

$$\beta_s = \begin{cases} \frac{\psi(X_s^x, \tilde{Z}_s^x) - \psi(X_s^x, Z_s^x)}{|\tilde{Z}_s^x - Z_s^x|^2} (\tilde{Z}_s^x - Z_s^x) & \text{if } \tilde{Z}_s^x - Z_s^x \neq 0 \\ 0 & \text{otherwise,} \end{cases},$$

and $\tilde{W}_t = -\int_0^t \beta_s ds + W_t$. According to the Girsanov theorem there exists a probability \mathbb{Q}_T under which $(\tilde{W}_t)_{t \in [0, T]}$ is a Brownian motion. Then we have

$$\bar{Y}_0^x = \underbrace{\mathbb{E}^{\mathbb{Q}_T} [\bar{Y}_T^x]}_{\leq M} + [\lambda - \tilde{\lambda}]T + [\mu - \tilde{\mu}] \underbrace{\mathbb{E}^{\mathbb{Q}_T} [K_T^x]}_{\geq 0}. \quad (2.16)$$

If we suppose that $\mu \leq \tilde{\mu}$ and $\lambda < \tilde{\lambda}$ then

$$\bar{Y}_0^x \leq [\lambda - \tilde{\lambda}]T + M \xrightarrow{T \rightarrow +\infty} -\infty;$$

this is a contradiction. So $\mu \leq \tilde{\mu} \Rightarrow \lambda \geq \tilde{\lambda}$. To show the continuity of λ we assume that $|\tilde{\mu} - \mu| \leq \varepsilon$ with $\varepsilon > 0$. Then

$$|\tilde{\lambda} - \lambda| = \frac{1}{T} \left| \mathbb{E}^{\mathbb{Q}_T} [\bar{Y}_0^x - \bar{Y}_T^x + [\tilde{\mu} - \mu]K_T^x] \right| \leq \frac{2M}{T} + \frac{\varepsilon}{T} \mathbb{E}^{\mathbb{Q}_T} [K_T^x].$$

Let us now prove a lemma about the bound on $\mathbb{E}^{\mathbb{Q}_T} [K_t^x]$.

Lemma 2.12 *There exists a constant C such that*

$$\mathbb{E}^{\mathbb{Q}_T} [K_t^x] \leq C(1+t), \quad \forall T \in \mathbb{R}^+, \forall t \in [0, T], \forall x \in \bar{G}.$$

Proof of the lemma. Applying Itô's formula to $\phi(X_t^x)$ we have for all $t \in \mathbb{R}^+$ and all $x \in \bar{G}$

$$K_t^x = \phi(X_t^x) - \phi(x) - \int_0^t \mathcal{L}\phi(X_s^x) ds - \int_0^t {}^t\nabla\phi(X_s^x)\sigma(X_s^x)dW_s. \quad (2.17)$$

Then

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}_T} [K_t^x] &= \mathbb{E}^{\mathbb{Q}_T} \left[\phi(X_t^x) - \phi(x) - \int_0^t \mathcal{L}\phi(X_s^x) ds - \int_0^t {}^t\nabla\phi(X_s^x)\sigma(X_s^x)(\beta_s ds + d\tilde{W}_s) \right] \\ &\leq \mathbb{E}^{\mathbb{Q}_T} \left[\underbrace{|\phi(X_t^x)|}_{\leq C/2} + \underbrace{|\phi(x)|}_{\leq C/2} + \int_0^t \underbrace{|\mathcal{L}\phi(X_s^x)|}_{\leq C/2} ds + \int_0^t \underbrace{|{}^t\nabla\phi(X_s^x)\sigma(X_s^x)\beta_s|}_{\leq C/2} ds \right] \\ &\leq C(1+t). \end{aligned}$$

□

Let us return back to the proof of Proposition 2.11. By applying Lemma 2.12 we obtain

$$|\tilde{\lambda} - \lambda| \leq \frac{2M}{T} + \frac{T+1}{T} C\varepsilon \xrightarrow{T \rightarrow +\infty} C\varepsilon.$$

The proof is therefore completed. □

To prove our second theorem of existence we need to introduce a further assumption.

(F2).

1. $|\psi|$ is bounded by M_ψ ;
2. $\mathbb{E}_\nu[\mathcal{L}\phi] := \int \mathcal{L}\phi d\nu < 0$ with ν the invariant measure for the process $(X_t)_{t \geq 0}$.

Theorem 2.13 (Existence of a solution) Assume that (G1), (G2), (G3), (H1), (H2), (H3), (F1) and (F2) hold true. Then for any $\lambda \in \mathbb{R}$ there exist $\mu \in \mathbb{R}$, $v \in \mathcal{C}_{lip}^0(\overline{G})$, a measurable function $\zeta : \mathbb{R}^d \rightarrow \mathbb{R}^{1 \times d}$ such that, if we define $Y_t^x := v(X_t^x)$ and $Z_t^x := \zeta(X_t^x)$ then $Z^x \in \mathcal{M}^2(\mathbb{R}^+, \mathbb{R}^{1 \times d})$ and \mathbb{P} -a.s. (Y^x, Z^x, μ) is a solution of EBSDE (2.13) with λ fixed, for all $x \in \overline{G}$. Moreover we have

$$|\lambda(\mu) - \lambda(0) - \mu \mathbb{E}_\nu[\mathcal{L}\phi]| \leq 2M_\psi.$$

Proof. Let $(Y, Z, \lambda(\mu))$ and $(\tilde{Y}, \tilde{Z}, \lambda(0))$ be two solutions of equation (2.13) linked to μ and 0 respectively. Let ν the invariant measure for the process $(X_t)_{t \geq 0}$. We set $\tilde{Y}^x := \tilde{Y}^x - Y^x$. Then, from equation (2.15), we deduce for all $T \in \mathbb{R}^+$

$$\int_{\overline{G}} \mathbb{E}[\tilde{Y}_0^x - \tilde{Y}_T^x - [\lambda(\mu) - \lambda(0)]T - \mu K_T^x] \nu(dx) = \int_{\overline{G}} \mathbb{E}\left[\int_0^T \left(\psi(X_s^x, \tilde{Z}_s^x) - \psi(X_s^x, Z_s^x)\right) ds\right] \nu(dx),$$

from which we obtain that

$$\left| \int_{\overline{G}} \mathbb{E}[\tilde{Y}_0^x - \tilde{Y}_T^x] \nu(dx) - [\lambda(\mu) - \lambda(0)]T - \mu \int_{\overline{G}} \mathbb{E}[K_T^x] \nu(dx) \right| \leq 2M_\psi T.$$

By using equation (2.17) we have

$$\begin{aligned} \int_{\overline{G}} \mathbb{E}[K_T^x] \nu(dx) &= \int_{\overline{G}} \mathbb{E}\left[\phi(X_T^x) - \phi(x) - \int_0^T \mathcal{L}\phi(X_s^x) ds\right] \nu(dx) \\ &= \mathbb{E}_\nu[\phi - \phi] - \int_0^T \int_{\overline{G}} \mathbb{E}[\mathcal{L}\phi(X_s^x)] \nu(dx) ds \\ &= -\mathbb{E}_\nu[\mathcal{L}\phi] T. \end{aligned}$$

Combining the last two relations, we get

$$\left| \frac{\int_{\overline{G}} \mathbb{E}[\tilde{Y}_0^x - \tilde{Y}_T^x] \nu(dx)}{T} - [\lambda(\mu) - \lambda(0)] + \mu \mathbb{E}_\nu[\mathcal{L}\phi] \right| \leq 2M_\psi.$$

Thus letting $T \rightarrow +\infty$ we conclude that

$$|\lambda(\mu) - \lambda(0) - \mu \mathbb{E}_\nu[\mathcal{L}\phi]| \leq 2M_\psi.$$

So, we obtain

$$\lambda(\mu) \xrightarrow{\mu \rightarrow +\infty} -\infty \quad \text{and} \quad \lambda(\mu) \xrightarrow{\mu \rightarrow -\infty} +\infty.$$

Finally the result is a direct consequence of the intermediate value theorem. \square

The hypothesis $\mathbb{E}_\nu[\mathcal{L}\phi] < 0$ says that the boundary has to be visited recurrently. When σ is non-singular on \overline{G} we show that this hypothesis is always verified.

Proposition 2.14 Assume that (G1), (G2) and (H1) hold true. We assume also that $\sigma(x)$ is non-singular for all $x \in \overline{G}$. Then for the invariant measure ν of the process $(X_t)_{t \geq 0}$ we have $\mathbb{E}_\nu[\mathcal{L}\phi] < 0$.

Proof. We already show that $\int_{\overline{G}} \mathbb{E}[K_T^x] \nu(dx) = -\mathbb{E}_\nu[\mathcal{L}\phi] T$, which implies that $\mathbb{E}_\nu[\mathcal{L}\phi] \leq 0$. If $\mathbb{E}_\nu[\mathcal{L}\phi] = 0$, then for ν -a.e. $x \in \overline{G}$ and for all $t \in \mathbb{R}^+$, $K_t^x = 0$. So, for ν -a.e. $x \in \overline{G}$, the process X^x is the solution of the stochastic differential equation

$$X_t^x = x + \int_0^t \tilde{b}(X_s^x) ds + \int_0^t \tilde{\sigma}(X_s^x) dW_s, \quad t \geq 0, \quad (2.18)$$

with \tilde{b} and $\tilde{\sigma}$ defined on \mathbb{R}^d by $\tilde{\sigma}(x) = \sigma(\text{proj}_{\overline{G}}(x))$ and $\tilde{b}(x) = b(\text{proj}_{\overline{G}}(x))$. But according to [53] (Corollary 2 of Theorem 7.1), the solution of equation (2.18) is a recurrent Markov process on \mathbb{R}^d . Thus this process is particularly unbounded: we have a contradiction. \square

When σ is singular on \overline{G} then (F2) is not necessarily verified.

Examples.

- Let $G = B(0, 1)$, $\phi(x) = \frac{1-|x|^2}{2}$, $b(x) = -x$ and $\sigma(x) = \begin{pmatrix} x_1 & 0 \\ & \ddots \\ 0 & x_d \end{pmatrix}$ on \overline{G} . Then δ_0 is the invariant measure and $\mathcal{L}\phi(0) = 0$. If we set $d = 1$, $\psi = 0$ and $g = 0$ then solutions of the differential equation (2.5) without boundary condition are $\{A_i + B_i x^3 - \frac{2}{3}\lambda \ln|x|, (A_i, B_i) \in \mathbb{R}^2\}$ on $[-1, 0[$ and $]0, 1]$. Thereby bounded continuous solutions are $\{A - \frac{\mu}{3}|x|^3, A \in \mathbb{R}\}$ and $\lambda(\mu) = 0$.
- Let $G = B(0, 1)$, $\phi(x) = \frac{1-|x|^2}{2}$, $b(x) = -x$ and $\sigma(x) = \begin{pmatrix} I_k & 0 \\ 0 & 0_{d-k} \end{pmatrix}$ on \overline{G} .
 $F_k := \{x \in \mathbb{R}^d / x_{k+1} = \dots = x_d = 0\} \simeq \mathbb{R}^k$ is a stationary subspace for solutions of equation (2.8).
Let ν_k the invariant measure on \mathbb{R}^k for $\tilde{\phi}(x) = \frac{1-|x|^2}{2}$, $\tilde{b}(x) = -x$ and $\tilde{\sigma}(x) = I_k$. According to Proposition 2.14, $\mathbb{E}_{\nu_k}[\tilde{\mathcal{L}}\tilde{\phi}] < 0$. Then $\nu := \nu_k \otimes \delta_{0_{\mathbb{R}^{d-k}}}$ is the invariant measure for the initial problem and $\mathbb{E}^\nu[\mathcal{L}\phi] < 0$.

Theorem 2.13 is not totally satisfactory for two reasons: we do not have a result on the uniqueness of μ and ψ is usually not bounded in optimal ergodic control problems. So we introduce another result of existence with a different hypothesis.

$$(F2'). \quad -\mathcal{L}\phi(x) > |{}^t\nabla\phi\sigma|_{\infty, \overline{G}} K_{\psi, z}, \quad \forall x \in \overline{G}.$$

Theorem 2.15 (Existence and uniqueness of a solution) Assume that (G1), (G2), (G3), (H1), (H2), (H3), (F1) and (F2') hold true. Then for any $\lambda \in \mathbb{R}$ there exist $\mu \in \mathbb{R}$, $v \in \mathcal{C}_{lip}^0(\overline{G})$, a measurable function $\zeta : \mathbb{R}^d \rightarrow \mathbb{R}^{1 \times d}$ such that, if we define $Y_t^x := v(X_t^x)$ and $Z_t^x := \zeta(X_t^x)$ then $Z^x \in \mathcal{M}^2(\mathbb{R}^+, \mathbb{R}^{1 \times d})$ and $\mathbb{P} - a.s.$ (Y^x, Z^x, μ) is a solution of EBSDE (2.13) with λ fixed, for all $x \in \overline{G}$. Moreover μ is unique among solutions (Y, Z, μ) with λ fixed such that Y is a bounded continuous adapted process and $Z \in \mathcal{M}^2(\mathbb{R}^+, \mathbb{R}^{1 \times d})$.

Proof. Let $(Y, Z, \lambda(\mu))$ and $(\tilde{Y}, \tilde{Z}, \lambda(\tilde{\mu}))$ be two solutions of equation (2.13) linked to μ and $\tilde{\mu}$. As in the proof of Proposition 2.11 we set $\bar{Y}^x := \tilde{Y}^x - Y^x$ and $\bar{Z}^x := \tilde{Z}^x - Z^x$. From equation (2.16), we have:

$$(\mu - \tilde{\mu})\mathbb{E}^{\mathbb{Q}^T} \left[\frac{K_T^x}{T} \right] = \frac{1}{T} \left(\bar{Y}_0^x - \mathbb{E}^{\mathbb{Q}^T} [\bar{Y}_T^x] \right) - (\lambda(\mu) - \lambda(\tilde{\mu})).$$

\bar{Y}^x is bounded, so $\mathbb{E}^{\mathbb{Q}^T} [K_T^x/T]$ has a limit $l_{\mu, \tilde{\mu}} \geq 0$ as $T \rightarrow +\infty$ when $\mu \neq \tilde{\mu}$ such that

$$(\lambda(\mu) - \lambda(\tilde{\mu})) + (\mu - \tilde{\mu})l_{\mu, \tilde{\mu}} = 0. \quad (2.19)$$

By the use of equation (2.17) we have

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}^T} \left[\frac{K_T^x}{T} \right] &= \mathbb{E}^{\mathbb{Q}^T} \left[\phi(X_T^x) - \phi(x) - \int_0^T \mathcal{L}\phi(X_s^x) ds - \int_0^T {}^t\nabla\phi(X_s^x) \sigma(X_s^x) \beta_s ds \right] \\ \mathbb{E}^{\mathbb{Q}^T} \left[\frac{K_T^x}{T} \right] &\geq -\frac{2|\phi|_\infty}{T} + \left[-\sup_{x \in \overline{G}} \mathcal{L}\phi - |{}^t\nabla\phi\sigma|_{\infty, \overline{G}} K_{\psi, z} \right]. \end{aligned}$$

We set $c = -\sup_{x \in \overline{G}} \mathcal{L}\phi - |{}^t\nabla\phi\sigma|_{\infty, \overline{G}} K_{\psi, z}$. Since hypothesis (F2') holds true, we have $c > 0$ and $l_{\mu, \tilde{\mu}} \geq c > 0$ when $\mu \neq \tilde{\mu}$. Thus, thanks to equation (2.19),

$$\lambda(\mu) \xrightarrow{\mu \rightarrow +\infty} -\infty \quad \text{and} \quad \lambda(\mu) \xrightarrow{\mu \rightarrow -\infty} +\infty.$$

Once again the existence result is a direct consequence of the intermediate value theorem. Moreover, if $\lambda(\mu) = \lambda(\tilde{\mu})$ then $\mu = \tilde{\mu}$. \square

Remark 2.16 By applying Lemma 2.12 we show that $\mathbb{E}^{\mathbb{Q}^T} [K_T^x/T]$ is bounded. So we have:

$$0 < c \leq l_{\mu, \tilde{\mu}} \leq C, \quad \forall \mu \neq \tilde{\mu}.$$

Remark 2.17 If we have interest in the second example dealt in this section we can see that (F2') hold true when $k/2 - 1 > K_{\psi, z}$.

2.4 Study of reflected Kolmogorov processes case

In this section, we assume that $((X_t)_{t \geq 0}, (\mathbb{P}_x)_{x \in \overline{G}})$ is a reflected Kolmogorov process with a unique invariant probability measure. We denote $\mathbb{P}_\nu(\cdot) := \int_{\overline{G}} \mathbb{P}_x(\cdot) \nu(dx)$. Then $((X_t)_{t \geq 0}, \mathbb{P}_\nu)$ is a stationary ergodic process. The aim is to obtain an equivalent to Theorem 2.15 with a less restrictive hypothesis than (F2'). We set $\sigma = \sqrt{2}I$ and $b = -\nabla U$ where $U : \mathbb{R}^d \rightarrow \mathbb{R}$ verifies the following assumptions:

(H4). $U \in \mathcal{C}^2(\mathbb{R}^d)$, ∇U is a Lipschitz function on \mathbb{R}^d and $\nabla^2 U \geq cI$ with $c > 0$.

We notice that (H4) implies (H3) and (H1). Moreover, without loss of generality, we use an extra assumption on ϕ :

(G4). $\nabla \phi$ is a Lipschitz function on \mathbb{R}^d .

To study the reflected process we will introduce the related penalized process:

$$X_t^{n,x} = x - \int_0^t \nabla U_n(X_s^{n,x}) ds + \sqrt{2} B_t, \quad t \geq 0, \quad x \in \mathbb{R}^d, \quad n \in \mathbb{N},$$

with $U_n = U + nd^2(\cdot, \overline{G})$. According to [45], $d^2(\cdot, \overline{G})$ is twice differentiable and $\nabla^2 d^2(\cdot, \overline{G}) \geq 0$. So, we have $\nabla^2 U_n \geq cI$. Let \mathcal{L}_n the transition semi-group generator of $(X_t^n)_{t \geq 0}$ with domain $\mathbb{D}_2(\mathcal{L}_n)$ and ν_n its invariant measure given by

$$\nu_n(dx) = \frac{1}{N_n} \exp(-U_n(x)) dx, \quad \text{with } N_n = \int_{\mathbb{R}^d} \exp(-U_n(x)) dx.$$

Proposition 2.18 $\mathbb{E}_{\nu_n}[f] \xrightarrow{n \rightarrow +\infty} \mathbb{E}_\nu[f]$ for all Lipschitz functions f .

The proof is given in the appendix. We obtain a simple corollary that we also prove in the appendix:

Corollary 2.19 $\nu(dx) = \frac{1}{N} \exp(-U(x)) 1_{x \in \overline{G}} dx$, with $N = \int_{\overline{G}} \exp(-U(x)) dx$. Moreover, ν_n converges narrowly to ν .

We now introduce a different assumption that will replace (F2'):

(F2''). $\left(\frac{\delta}{\sqrt{2c}} + \sqrt{2} |\nabla \phi|_{\infty, \overline{G}} \right) K_{\psi,z} < -\mathbb{E}_\nu[\mathcal{L}\phi]$,
with $\delta = \sup_{x \in \overline{G}} ({}^t \nabla U(x)x) - \inf_{x \in \overline{G}} ({}^t \nabla U(x)x)$.

Theorem 2.20 (Existence and uniqueness of a solution) Theorem 2.15 remains true if we assume that (G1), (G2), (G3), (G4), (H2), (H4), (F1) and (F2'') hold.

Proof. If we use notations of the previous section, it is sufficient to show that there exists a constant $C > 0$ such that $\lim_{T \rightarrow +\infty} \int_{\overline{G}} \mathbb{E}^{\mathbb{Q}^T} \left[\frac{K_{\psi,z}}{T} \right] \nu(dx) \geq C$ for all $\mu \neq \tilde{\mu}$. Thanks to hypothesis (F2'') we pick ε in the interval

$$\left[\frac{\delta}{\sqrt{2c}} K_{\psi,z}, -\mathbb{E}_\nu[\mathcal{L}\phi] - \sqrt{2} |\nabla \phi|_{\infty, \overline{G}} K_{\psi,z} \right],$$

and consider the set

$$A_T := \left\{ -\frac{1}{T} \int_0^T \mathcal{L}\phi(X_s) ds \leq -\mathbb{E}_\nu[\mathcal{L}\phi] - \varepsilon \right\} \subset \overline{G} \times \Omega,$$

with $T > 0$. We have

$$\begin{aligned}
\int_{\overline{G}} \mathbb{E}^{\mathbb{Q}_T} \left[\frac{K_T^x}{T} \right] \nu(dx) &= \int_{\overline{G}} \mathbb{E}^{\mathbb{Q}_T} \left[\frac{\phi(X_T^x)}{T} - \frac{\phi(x)}{T} - \frac{1}{T} \int_0^T \mathcal{L}\phi(X_s^x) ds \right. \\
&\quad \left. - \frac{\sqrt{2}}{T} \int_0^T {}^t\nabla\phi(X_s^x) \beta_s ds \right] \nu(dx) \\
&\geq -\frac{2|\phi|_\infty}{T} + \int_{\overline{G}} \mathbb{E}^{\mathbb{Q}_T} \left[(-\mathbb{E}_\nu[\mathcal{L}\phi] - \varepsilon) 1_{c_{A_T}} - |\mathcal{L}\phi|_{\infty, \overline{G}} 1_{A_T} \right] \nu(dx) \\
&\quad - \sqrt{2} |\nabla\phi|_{\infty, \overline{G}} K_{\psi, z} \\
&\geq -\frac{2|\phi|_\infty}{T} + (-\mathbb{E}_\nu[\mathcal{L}\phi] - \varepsilon) \left(1 - \int_{\overline{G}} \mathbb{Q}_T(A_T) \nu(dx) \right) \\
&\quad - |\mathcal{L}\phi|_{\infty, \overline{G}} \int_{\overline{G}} \mathbb{Q}_T(A_T) \nu(dx) - \sqrt{2} |\nabla\phi|_{\infty, \overline{G}} K_{\psi, z}.
\end{aligned}$$

By using Hölder's inequality with $p > 1$ and $q > 1$ such that $1/p + 1/q = 1$ we obtain, for all $x \in \overline{G}$,

$$\begin{aligned}
\mathbb{Q}_T(A_T) &= \mathbb{E} \left[\exp \left(\int_0^T \beta_s dW_s - \frac{1}{2} \int_0^T |\beta_s|^2 ds \right) 1_{A_T} \right] \\
&\leq \mathbb{E} \left[\exp \left(p \int_0^T \beta_s dW_s - \frac{p^2}{2} \int_0^T |\beta_s|^2 ds + \frac{p(p-1)}{2} \int_0^T |\beta_s|^2 ds \right) \right]^{1/p} \mathbb{P}_x(A_T)^{1/q} \\
&\leq \exp \left(\frac{(p-1)}{2} K_{\psi, z}^2 T \right) \mathbb{P}_x(A_T)^{1-1/p}.
\end{aligned}$$

So

$$\int_{\overline{G}} \mathbb{Q}_T(A_T) \nu(dx) \leq \exp \left(\frac{(p-1)}{2} K_{\psi, z}^2 T \right) \mathbb{P}_\nu(A_T)^{1-1/p}.$$

To conclude we are going to use the following proposition which will be proved in the appendix thanks to Theorem 3.1 of [52]:

Proposition 2.21 *Assume that (G1), (G2), (G3), (G4), (H1) and (H4) hold. Then*

$$\mathbb{P}_\nu(A_T) \leq \exp \left(-\frac{c\varepsilon^2 T}{\delta^2} \right).$$

So

$$\int_{\overline{G}} \mathbb{Q}_T(A_T) \nu(dx) \leq \exp \left[\underbrace{\left(\frac{p(p-1)}{2} K_{\psi, z}^2 - \frac{(p-1)c\varepsilon^2}{\delta^2} \right)}_{B_p} \frac{T}{p} \right].$$

B_p is a trinomial in p that has two different real roots 1 and $\frac{2c\varepsilon^2}{\delta^2 K_{\psi, z}^2} > 1$ because $\varepsilon > \delta K_{\psi, z} / \sqrt{2c}$ by hypothesis (F2''). So we are able to find $p > 1$ such that $B_p < 0$. Then $\int_{\overline{G}} \mathbb{Q}_T(A_T) \nu(dx) \xrightarrow{T \rightarrow +\infty} 0$ and

$$\lim_{T \rightarrow +\infty} \int_{\overline{G}} \mathbb{E}^{\mathbb{Q}_T} \left[\frac{K_T^x}{T} \right] \nu(dx) \geq -\mathbb{E}_\nu[\mathcal{L}\phi] - \sqrt{2} |\nabla\phi|_{\infty, \overline{G}} K_{\psi, z} - \varepsilon > 0.$$

□

Remark 2.22 *All these results stay true if $\sigma(x) = \sqrt{2} \begin{pmatrix} I_k & 0 \\ 0 & 0_{d-k} \end{pmatrix}$ and F_k , defined in the previous example, is a stationary subspace of ∇U . We can even replace (F2'') by*

$$\left(\sqrt{\frac{1}{2c}} \delta + \sqrt{2} |\nabla\phi|_{\infty, \overline{G} \cap F_k} \right) K_{\psi, z} < -\mathbb{E}_\nu[\mathcal{L}\phi],$$

with $\delta = \sup_{x \in \overline{G} \cap F_k} ({}^t\nabla U(x)x) - \inf_{x \in \overline{G} \cap F_k} ({}^t\nabla U(x)x)$. Indeed, as we see in the previous example, ν is non-zero at most on the set $\overline{G} \cap F_k$. So it is possible to restrict the process to the subspace F_k .

2.5 Probabilistic interpretation of the solution of an elliptic PDE with linear Neumann boundary condition

Consider the semi-linear elliptic PDE:

$$\begin{cases} \mathcal{L}v(x) + \psi(x, {}^t \nabla v(x) \sigma(x)) = \lambda, & x \in G \\ \frac{\partial v}{\partial n}(x) + g(x) = \mu, & x \in \partial G, \end{cases} \quad (2.20)$$

with

$$\mathcal{L}f(x) = \frac{1}{2} \text{Tr}(\sigma(x) {}^t \sigma(x) \nabla^2 f(x)) + {}^t b(x) \nabla f(x).$$

The solution v of this semi-linear problem is a scalar function and λ, μ are constants: In some cases μ is a given constant and λ is part of the unknowns, and it is the contrary in the other cases. We will see now that v , defined in Theorem 2.9, in Theorem 2.13, in Theorem 2.15 or in Theorem 2.20, is a viscosity solution of PDE (2.20). See e.g. [75] Definition 5.2 for the definition of a viscosity solution.

Theorem 2.23 *Assume hypothesis of Theorem 2.9. For any $\mu \in \mathbb{R}$, there exists a solution $(v(X.), \zeta(X.), \lambda)$ of the EBSDE (2.9) with μ fixed. Then $(v, \lambda) \in \mathcal{C}_{lip}^0(\overline{G}) \times \mathbb{R}$ is a viscosity solution of the elliptic PDE (2.20) with μ fixed.*

Theorem 2.24 *Assume hypothesis of Theorem 2.13, Theorem 2.15 or Theorem 2.20. For any $\lambda \in \mathbb{R}$, there exists a solution $(v(X.), \zeta(X.), \mu)$ of the EBSDE (2.9) with λ fixed. Then $(v, \mu) \in \mathcal{C}_{lip}^0(\overline{G}) \times \mathbb{R}$ is a viscosity solution of the elliptic PDE (2.20) with λ fixed.*

The proof of these results is very standard and can be easily adapted from [75], Theorem 4.3.

Remark 2.25 *For any λ , uniqueness of solution (v, μ) , up to additive constants for v , is given by Barles and Da Lio in Theorem 4.4 of [11] with different hypothesis:*

- G is a bounded domain with a $W^{3,\infty}$ boundary,
- the outward unit normal vector to ∂G is a Lipschitz continuous function on ∂G ,
- there exists $K \geq 0$ such that

$$|\psi(x, z) - \psi(x', z')| \leq K(|x - x'| (1 + |z| + |z'|) + |z - z'|), \quad \forall x, x' \in \overline{G}, z, z' \in \mathbb{R}^d,$$

- σ is non-singular on \overline{G} ,
- there exists a continuous function ψ_∞ such that

$$t^{-1} \psi(x, tz) \rightarrow \psi_\infty(x, z) \text{ locally uniformly, as } t \rightarrow +\infty,$$

- g is a Lipschitz continuous function on \overline{G} .

If σ is non-singular on \overline{G} we notice that it is possible to jointly modify b and ψ without modifying PDE 2.20. We set $\tilde{b}(x) = b(x) - \xi x$ and $\tilde{\psi}(x, z) = \psi(x, z) + \xi z \sigma^{-1}(x) x$ for $\xi \in \mathbb{R}^+$. Then we are able to find a new hypothesis substituting (H3). We note $\tilde{\eta}$ the scalar η corresponding to \tilde{b} .

Proposition 2.26 *If $\eta + K_{\psi,z} K_\sigma < 0$ or $K_\sigma \sup_{x \in \overline{G}} |\sigma^{-1}(x) x| < 1$ then there exists $\xi \geq 0$ such that $\tilde{\eta} + K_{\tilde{\psi},z} K_\sigma < 0$. In particular it is true when σ is a constant function.*

Proof: It suffices to notice that $\tilde{\eta} = \eta - \xi$ and $K_{\tilde{\psi},z} \leq K_{\psi,z} + \xi \sup_{x \in \overline{G}} |\sigma^{-1}(x) x|$. So

$$\tilde{\eta} + K_{\tilde{\psi},z} K_\sigma \leq \eta + K_{\psi,z} K_\sigma + \xi (K_\sigma \sup_{x \in \overline{G}} |\sigma^{-1}(x) x| - 1).$$

□

2.6 Optimal ergodic control

Let U be a separable metric space. We define a control ρ as an (\mathcal{F}_t) -progressively measurable U valued process. We introduce $R : U \rightarrow \mathbb{R}^d$ and $L : \mathbb{R}^d \times U \rightarrow \mathbb{R}$ two continuous functions such that, for some constants $M_R > 0$ and $M_L > 0$,

$$|R(u)| \leq M_R, \quad |L(x, u)| \leq M_L, \quad |L(x, u) - L(x', u)| \leq c|x - x'|, \quad \forall u \in U, x, x' \in \mathbb{R}^d. \quad (2.21)$$

Given an arbitrary control ρ and $T > 0$, we introduce the Girsanov density

$$\Gamma_T^\rho = \exp \left(\int_0^T R(\rho_s) dW_s - \frac{1}{2} \int_0^T |R(\rho_s)|^2 ds \right)$$

and the probability $d\mathbb{P}_T^\rho = \Gamma_T^\rho d\mathbb{P}$ on \mathcal{F}_T . Ergodic costs corresponding to a given control ρ and a starting point $x \in \mathbb{R}^d$ are defined in the following way:

$$I(x, \rho) = \limsup_{T \rightarrow +\infty} \frac{1}{T} \mathbb{E}^{\rho, T} \left[\int_0^T L(X_s^x, \rho_s) ds + \int_0^T [g(X_s^x) - \mu] dK_s^x \right], \quad (2.22)$$

$$J(x, \rho) = \limsup_{T \rightarrow +\infty} \begin{cases} +\infty & \text{if } \mathbb{E}^{\rho, T}[K_T^x] = 0, \\ \frac{1}{\mathbb{E}^{\rho, T}[K_T^x]} \mathbb{E}^{\rho, T} \left[\int_0^T [L(X_s^x, \rho_s) - \lambda] ds + \int_0^T g(X_s^x) dK_s^x \right] & \text{otherwise,} \end{cases} \quad (2.23)$$

where $\mathbb{E}^{\rho, T}$ denotes expectation with respect to \mathbb{P}_T^ρ . We notice that $W_t^\rho = W_t - \int_0^t R(\rho_s) ds$ is a Wiener process on $[0, T]$ under \mathbb{P}_T^ρ .

Our purpose is to minimize costs I and J over all controls. So we first define the Hamiltonian in the usual way

$$\psi(x, z) = \inf_{u \in U} \{L(x, u) + zR(u)\}, \quad x \in \mathbb{R}^d, z \in \mathbb{R}^{1 \times d}, \quad (2.24)$$

and we remark that if, for all x, z , the infimum is attained in (2.24) then, according to Theorem 4 of [68], there exists a measurable function $\gamma : \mathbb{R}^d \times \mathbb{R}^{1 \times d} \rightarrow U$ such that

$$\psi(x, z) = L(x, \gamma(x, z)) + zR(\gamma(x, z)).$$

We notice that ψ is a Lipschitz function: hypothesis (H2) is verified with $K_{\psi, z} = M_R$.

Theorem 2.27 *Assume that hypotheses of Theorem 2.9 holds true. Let (Y, Z, λ) be a solution of (2.13) with μ fixed. Then the following hold:*

1. *For arbitrary control ρ we have $I(x, \rho) \geq \lambda$ and the equality holds if $L(X_t^x, \rho_t) + Z_t^x R(\rho_t) = \psi(X_t^x, Z_t^x)$, \mathbb{P} -a.s. for almost every t .*
2. *If the infimum is attained in (2.24) then the control $\bar{\rho}_t = \gamma(X_t^x, Z_t^x)$ verifies $I(x, \bar{\rho}) = \lambda$.*

This theorem can be proved in the same manner as that of Theorem 7.1 in [43], so we omit it.

Remark 2.28 1. *If the infimum is attained in (2.24) then there exists an optimal feedback control given by the function $x \mapsto \gamma(x, \xi(x))$ where $(Y, \xi(X), \lambda)$ is the solution constructed in Theorem 2.9.*

2. *If limsup is changed into liminf in the definition (2.22) of the cost, then the same conclusion holds, with the obvious modifications, and the optimal value is given by λ in both cases.*

Theorem 2.29 *Assume that hypotheses of Theorem 2.15 or Theorem 2.20 hold true. Let (Y, Z, μ) a solution of (2.13) with λ fixed. Then the following hold:*

1. *For arbitrary control ρ we have $J(x, \rho) \geq \mu$ and the equality holds if $L(X_t^x, \rho_t) + Z_t^x R(\rho_t) = \psi(X_t^x, Z_t^x)$, \mathbb{P} -a.s. for almost every t .*
2. *If the infimum is attained in (2.24) then the control $\bar{\rho}_t = \gamma(X_t^x, Z_t^x)$ verifies $J(x, \bar{\rho}) = \mu$.*

Proof. As (Y, Z, μ) is a solution of the EBSDE with λ fixed, we have

$$\begin{aligned} -dY_t^x &= [\psi(X_t^x, Z_t^x) - \lambda]dt + [g(X_t^x) - \mu]dK_t^x - Z_t^x dW_t \\ &= [\psi(X_t^x, Z_t^x) - \lambda]dt + [g(X_t^x) - \mu]dK_t^x - Z_t^x dW_t^\rho - Z_t^x R(\rho_t)dt, \end{aligned}$$

from which we deduce that

$$\begin{aligned} \mu \mathbb{E}^{\rho, T}[K_T^x] &= \mathbb{E}^{\rho, T}[Y_T^x - Y_0^x] + \mathbb{E}^{\rho, T} \left[\int_0^T [\psi(X_t^x, Z_t^x) - Z_t^x R(\rho_t) - L(X_t^x, \rho_t)]dt \right] \\ &\quad + \mathbb{E}^{\rho, T} \left[\int_0^T [L(X_t^x, \rho_t) - \lambda]dt \right] + \mathbb{E}^{\rho, T} \left[\int_0^T g(X_t^x) dK_t^x \right]. \end{aligned}$$

Thus

$$\mu \mathbb{E}^{\rho, T}[K_T^x] + \mathbb{E}^{\rho, T}[Y_0^x - Y_T^x] \leq \mathbb{E}^{\rho, T} \left[\int_0^T [L(X_t^x, \rho_t) - \lambda]dt + \int_0^T g(X_t^x) dK_t^x \right].$$

To conclude we are going to use the following lemma that we will prove immediately after the proof of this theorem:

Lemma 2.30 *Assume that hypotheses of Theorem 2.15 or Theorem 2.20 hold true. Then for all $x \in \overline{G}$*

$$\lim_{T \rightarrow +\infty} \mathbb{E}^{\rho, T}[K_T^x] = +\infty.$$

So, for $T > T_0$, $\mathbb{E}^{\rho, T}[K_T^x] > 0$ and

$$\mu + \frac{\mathbb{E}^{\rho, T}[Y_0^x - Y_T^x]}{\mathbb{E}^{\rho, T}[K_T^x]} \leq \frac{1}{\mathbb{E}^{\rho, T}[K_T^x]} \mathbb{E}^{\rho, T} \left[\int_0^T [L(X_t^x, \rho_t) - \lambda]dt + \int_0^T g(X_t^x) dK_t^x \right].$$

Since Y is bounded we finally obtain

$$\mu \leq \limsup_{T \rightarrow +\infty} \frac{1}{\mathbb{E}^{\rho, T}[K_T^x]} \mathbb{E}^{\rho, T} \left[\int_0^T [L(X_t^x, \rho_t) - \lambda]dt + \int_0^T g(X_t^x) dK_t^x \right] = J(x, \rho).$$

Similarly, if $L(X_t^x, \rho_t) + Z_t^x R(\rho_t) = \psi(X_t^x, Z_t^x)$,

$$\mu \mathbb{E}^{\rho, T}[K_T^x] + \mathbb{E}^{\rho, T}[Y_0^x - Y_T^x] = \mathbb{E}^{\rho, T} \left[\int_0^T [L(X_t^x, \rho_t) - \lambda]dt + \int_0^T g(X_t^x) dK_t^x \right],$$

and the claim holds. \square

Proof of Lemma 2.30. Firstly we assume that hypotheses of Theorem 2.15 hold true. As in the proof of this theorem, we have by using equation (2.17),

$$\mathbb{E}^{\rho, T}[K_T^x] = \mathbb{E}^{\rho, T} \left[\phi(X_T^x) - \phi(x) - \int_0^T \mathcal{L}\phi(X_s^x)ds - \int_0^T {}^t\nabla\phi(X_s^x)\sigma(X_s^x)R(\rho_s)ds \right],$$

from which we deduce that

$$\mathbb{E}^{\rho, T} \left[\frac{K_T^x}{T} \right] \geq -\frac{2|\phi|_\infty}{T} + \left[-\sup_{x \in \overline{G}} \mathcal{L}\phi(x) - |\nabla\phi\sigma|_{\infty, \overline{G}} M_R \right].$$

Thanks to hypothesis (F2') we have

$$\mathbb{E}^{\rho, T} \left[\frac{K_T^x}{T} \right] \geq \frac{1}{2} \left[-\sup_{x \in \overline{G}} \mathcal{L}\phi(x) - |\nabla\phi\sigma|_{\infty, \overline{G}} M_R \right] > 0, \quad \forall T > T_0,$$

and the claim is proved. We now assume that hypotheses of Theorem 2.20 hold true. Let ν be the invariant measure of $(X_t)_{t \geq 0}$. Exactly as in the proof of Theorem 2.20 we are able to show that

$$\int_{\overline{G}} \mathbb{E}^{\rho, T} [K_T^y / T] \nu(dy) \geq C > 0$$

for all $T > T_0$ by replacing β with $R(\rho)$. On the other hand, for all $x \in \overline{G}$ and $T \in \mathbb{R}^{+*}$, we have

$$\left| \frac{\int_{\overline{G}} \mathbb{E}^{\rho, T} [K_T^y] \nu(dy) - \mathbb{E}^{\rho, T} [K_T^x]}{T} \right| \leq \frac{1}{T} \int_{\overline{G}} \mathbb{E}^{\rho, T} [|K_T^y - K_T^x|] \nu(dy)$$

and

$$\begin{aligned} \frac{1}{T} \int_{\overline{G}} \mathbb{E}^{\rho, T} [|K_T^y - K_T^x|] \nu(dy) &\leq \frac{4|\phi|_\infty}{T} + \frac{1}{T} \int_{\overline{G}} \mathbb{E}^{\rho, T} \left[\int_0^T |\mathcal{L}\phi(X_s^y) - \mathcal{L}\phi(X_s^x)| ds \right] \nu(dy) \\ &\quad + \frac{1}{T} \int_{\overline{G}} \mathbb{E}^{\rho, T} \left[\int_0^T |{}^t\nabla\phi(X_s^y)\sigma(X_s^y) - {}^t\nabla\phi(X_s^x)\sigma(X_s^x)| |R(\rho_s)| ds \right] \nu(dy). \end{aligned}$$

Since $\mathcal{L}\phi$ and ${}^t\nabla\phi\sigma$ are Lipschitz functions, we obtain

$$\begin{aligned} \left| \frac{\int_{\overline{G}} \mathbb{E}^{\rho, T} [K_T^y] \nu(dy) - \mathbb{E}^{\rho, T} [K_T^x]}{T} \right| &\leq \frac{4|\phi|_\infty}{T} + \frac{K_{\mathcal{L}\phi}}{T} \int_{\overline{G}} \mathbb{E}^{\rho, T} \left[\int_0^T |X_s^y - X_s^x| ds \right] \nu(dy) \\ &\quad + \frac{M_R K_{{}^t\nabla\phi\sigma}}{T} \int_{\overline{G}} \mathbb{E}^{\rho, T} \left[\int_0^T |X_s^y - X_s^x| ds \right] \nu(dy). \end{aligned}$$

Exactly as in Lemma 2.5 we are able to show that for all $s \geq 0$

$$\mathbb{E}^{\rho, T} [|X_s^y - X_s^x|^2] \leq e^{2(\eta + M_R K_\sigma)s} |y - x|^2.$$

Finally,

$$\begin{aligned} \left| \frac{\int_{\overline{G}} \mathbb{E}^{\rho, T} [K_T^y] \nu(dy) - \mathbb{E}^{\rho, T} [K_T^x]}{T} \right| &\leq \frac{K_{\mathcal{L}\phi} + M_R K_{{}^t\nabla\phi\sigma}}{T} \int_{\overline{G}} |y - x|^2 \nu(dy) \int_0^T e^{(\eta + M_R K_\sigma)s} ds \\ &\quad + \frac{4|\phi|_\infty}{T} \\ &\leq \frac{K_{\mathcal{L}\phi} + M_R K_{{}^t\nabla\phi\sigma}}{T} \int_{\overline{G}} |y - x|^2 \nu(dy) \frac{1 - e^{(\eta + M_R K_\sigma)T}}{-\eta - M_R K_\sigma} \\ &\quad + \frac{4|\phi|_\infty}{T}. \end{aligned}$$

Since hypothesis (H3) holds true, $\eta + M_R K_\sigma < 0$ and so

$$\lim_{T \rightarrow +\infty} \left| \frac{\int_{\overline{G}} \mathbb{E}^{\rho, T} [K_T^y] \nu(dy) - \mathbb{E}^{\rho, T} [K_T^x]}{T} \right| = 0.$$

Thus, for all $x \in \overline{G}$ there exists $T_0 > 0$ such that

$$\mathbb{E}^{\rho, T} [K_T^x / T] \geq \frac{1}{2} \int_{\overline{G}} \mathbb{E}^{\rho, T} [K_T^y / T] \nu(dy) \geq c/2 > 0, \quad \forall T \geq T_0,$$

and the claim follows. \square

Remark 2.31 Remark 2.28 remains true for Theorem 2.29.

2.7 Some additional results: EBSDEs on a non-convex bounded set

In the previous sections we have supposed that G was a bounded convex set. We shall substitute hypothesis (G2) by this one:

(G2'). G is a bounded subset of \mathbb{R}^d .

In this section we suppose also that σ is a constant function. At last, we set

$$\alpha = \sup_{x \in co(\bar{G})} \sup_{|y|=1} ({}^t y \nabla^2 \phi(x) y)$$

with $co(\bar{G})$ the convex hull of \bar{G} . Without loss of generality we assume that $\alpha > 0$. Indeed, $\alpha \leq 0$ if and only if ϕ is concave which implies that \bar{G} is a convex set. In the previous sections the convexity of G has been used to prove Lemma 2.5 so we will modify it:

Lemma 2.32 *Assume (G1), (G2'), (H1), (H2) hold true and σ is a constant function. Let*

$$\begin{aligned} \theta := & \sup_{x, y \in \bar{G}, x \neq y, z, z' \in \mathbb{R}^d, z \neq z'} \left\{ 2 \frac{{}^t(x-y)(b(x)-b(y))}{|x-y|^2} \right. \\ & - \alpha {}^t(\nabla \phi(x) + \nabla \phi(y)) \sigma \beta(x, y, z, z') \\ & - \frac{\alpha}{2} \text{Tr}(\nabla^2 \phi(x) \sigma^t \sigma + \nabla^2 \phi(y) \sigma^t \sigma) - \alpha {}^t \nabla \phi(x) b(x) - \alpha {}^t \nabla \phi(y) b(y) \\ & \left. + \alpha^2 ({}^t \nabla \phi(x) + {}^t \nabla \phi(y)) \sigma^t \sigma (\nabla \phi(x) + \nabla \phi(y)) \right\}, \end{aligned}$$

with

$$\beta(x, y, z, z') = \begin{cases} \frac{\psi(x, z') + \psi(y, z') - \psi(y, z) - \psi(x, z)}{2|z' - z|^2} {}^t(z' - z) & \text{if } z \neq z' \\ 0 & \text{otherwise.} \end{cases}$$

Then there exists a constant M which depends only on ϕ and such that for all $0 \leq t \leq s \leq T$,

$$\mathbb{E}^{\mathbb{Q}^T} \left[|X_s^x - X_s^{x'}|^2 \middle| \mathcal{F}_t \right] \leq M e^{\theta(s-t)} |X_t^x - X_t^{x'}|^2.$$

Proof. Firstly we show an elementary lemma.

Lemma 2.33 $\forall x \in \bar{G}, \forall y \in \partial G$ we have

$$-\alpha |x - y|^2 + 2 {}^t(y - x) \nabla \phi(y) \leq 0.$$

Proof. Let $x \in \bar{G}$ and $y \in \partial G$. According to Taylor-Lagrange theorem there exists $t \in]0, 1[$ such that

$$\phi(x) = \phi(y) + {}^t(x - y) \nabla \phi(y) + \frac{1}{2} {}^t(x - y) \nabla^2 \phi(tx + (1 - t)y)(x - y).$$

$\phi(x) \geq 0$, $\phi(y) = 0$ and the claim easily follows. □

As in Lions and Sznitman [65] page 524, using Itô's formula, we develop the semi-martingale

$$e^{-\theta u} e^{-\alpha(\phi(X_u^x) + \phi(X_u^{x'}))} |X_u^x - X_u^{x'}|^2,$$

which leads us to

$$\begin{aligned}
d\left(e^{-\theta u}e^{-\alpha(\phi(X_u^x)+\phi(X_u^{x'}))}|X_u^x - X_u^{x'}|^2\right) = & \\
& -\theta e^{-\theta u}e^{-\alpha(\phi(X_u^x)+\phi(X_u^{x'}))}|X_u^x - X_u^{x'}|^2 du \\
& + 2e^{-\theta u}e^{-\alpha(\phi(X_u^x)+\phi(X_u^{x'}))}\left[{}^t(X_u^x - X_u^{x'})(b(X_u^x) - b(X_u^{x'}))du\right. \\
& \quad \left.+ {}^t(X_u^x - X_u^{x'})\nabla\phi(X_u^x)dK_u^x - {}^t(X_u^x - X_u^{x'})\nabla\phi(X_u^{x'})dK_u^{x'}\right] \\
& - \alpha e^{-\theta u}e^{-\alpha(\phi(X_u^x)+\phi(X_u^{x'}))}|X_u^x - X_u^{x'}|^2\left[dK_u^x + dK_u^{x'}\right. \\
& \quad \left.+ {}^t(\nabla\phi(X_u^x) + \nabla\phi(X_u^{x'}))\sigma(d\tilde{W}_u + \beta_u du)\right. \\
& \quad \left.+ \frac{1}{2}\text{Tr}(\nabla^2\phi(X_u^x)\sigma^t\sigma + \nabla^2\phi(X_u^{x'})\sigma^t\sigma)du\right. \\
& \quad \left.+ ({}^t\nabla\phi(X_u^x)b(X_u^x) + {}^t\nabla\phi(X_u^{x'})b(X_u^{x'}))du\right] \\
& + \alpha^2 e^{-\theta u}e^{-\alpha(\phi(X_u^x)+\phi(X_u^{x'}))}|X_u^x - X_u^{x'}|^2 \\
& \quad \left[{}^t(\nabla\phi(X_u^x) + \nabla\phi(X_u^{x'}))\sigma^t\sigma(\nabla\phi(X_u^x) + \nabla\phi(X_u^{x'}))\right]du.
\end{aligned}$$

By Lemma 2.33 we have

$$\left(2{}^t(X_u^x - X_u^{x'})\nabla\phi(X_u^x) - \alpha|X_u^x - X_u^{x'}|^2\right)dK_u^x \leq 0,$$

and

$$\left(2{}^t(X_u^{x'} - X_u^x)\nabla\phi(X_u^{x'}) - \alpha|X_u^x - X_u^{x'}|^2\right)dK_u^{x'} \leq 0.$$

Applying the definitions of β and θ , we obtain

$$\begin{aligned}
d\left(e^{-\theta u}e^{-\alpha(\phi(X_u^x)+\phi(X_u^{x'}))}|X_u^x - X_u^{x'}|^2\right) \leq & \\
& -\alpha e^{-\alpha(\phi(X_u^x)+\phi(X_u^{x'}))}|X_u^x - X_u^{x'}|^2({}^t\nabla\phi(X_u^x) + \nabla\phi(X_u^{x'}))\sigma d\tilde{W}_u.
\end{aligned}$$

Thereby, for all $0 \leq t \leq s \leq n$

$$\mathbb{E}^{\mathbb{Q}^T}\left[e^{-\theta(s-t)-\alpha(\phi(X_s^x)+\phi(X_s^{x'}))}|X_s^x - X_s^{x'}|\middle|\mathcal{F}_t\right] \leq |X_t^x - X_t^{x'}|.$$

The claim follows by setting $M = e^{2\alpha \sup_{x \in \bar{G}} \phi(x)}$. □

We replace hypothesis (H3) by an analogous one:

(H3'). $\theta < 0$.

In the context of this section, Lemma 2.32 plays the same role as Lemma 2.5 in section 2.2. So, with the new hypothesis (H3'), we are able to show an analogous of Theorem 2.6 in a similar way. Moreover, since Theorems 2.9, 2.13 and 2.15 are consequences of Theorem 2.6, we are able to do the same thing to these ones.

Theorem 2.34 *Assume that σ is a constant function. Theorems 2.6, 2.9, 2.13 and 2.15 stay true if we substitute hypothesis (G2) and (H3) by (G2') and (H3').*

As in section 2.5, it is possible to jointly modify b and ψ without modifying the PDE 2.20 if σ is non-singular on \bar{G} . We set $\tilde{b}(x) = b(x) - \xi x$ and $\tilde{\psi}(x, z) = \psi(x, z) + \xi z \sigma^{-1}x$ for $\xi \in \mathbb{R}^+$. Then we are able to find a new hypothesis substituting (H3'). We note $\theta(\xi)$ the scalar θ corresponding to \tilde{b} and $\tilde{\psi}$. Let d be the diameter of \bar{G} :

$$d := \sup_{x, y \in \bar{G}} |x - y|.$$

Proposition 2.35 $\tilde{\theta}(\xi) \leq \theta - (2 - \frac{1}{2}d^2\alpha^2)\xi$. *In particular, if $\alpha d < 2$ then there exists $\xi \geq 0$ such that $\tilde{\theta}(\xi) < 0$.*

Proof. Let $\tilde{\beta}$ be the function β linked with $\tilde{\psi}$. We have $\tilde{\beta}(x, y, z, z') = \beta(x, y, z, z') + \frac{\xi}{2}\sigma^{-1}(x + y)$. Thus $\tilde{\theta}(\xi) \leq \theta + C\xi$ with

$$\begin{aligned} C &= -2 + \sup_{x, y \in \bar{G}, x \neq y} \left\{ -\frac{\alpha}{2} {}^t(\nabla\phi(x) + \nabla\phi(y))(x + y) + \alpha({}^t\nabla\phi(x)x + {}^t\nabla\phi(y)y) \right\} \\ &= -2 + \frac{\alpha}{2} \sup_{x, y \in \bar{G}} \left\{ {}^t(\nabla\phi(x) - \nabla\phi(y))(x - y) \right\}. \end{aligned}$$

On the other hand, we have

$$\sup_{x, y \in \bar{G}} \left\{ {}^t(\nabla\phi(x) - \nabla\phi(y))(x - y) \right\} \leq d^2\alpha.$$

Indeed, according to the Taylor Lagrange theorem there exist $t, t' \in]0, 1[$ such that

$$\phi(x) = \phi(y) + {}^t(x - y)\nabla\phi(y) + \frac{1}{2} {}^t(x - y)\nabla^2\phi(tx + (1 - t)y)(x - y),$$

$$\phi(y) = \phi(x) + {}^t(y - x)\nabla\phi(x) + \frac{1}{2} {}^t(y - x)\nabla^2\phi(t'y + (1 - t')x)(y - x).$$

Finally $C \leq -2 + \frac{d^2\alpha^2}{2}$ and the proof is therefore completed. \square

The assumption obtained in Proposition 2.35 is very simple but it could be not optimal. Indeed, α strongly depends on ϕ and it could be huge if the function ϕ is not well-chosen. So, we are able to optimize the choice of ϕ by defining

$$\alpha^* = \inf_{\phi \in \mathcal{A}} \left[\sup_{x \in \text{co}(\bar{G})} \sup_{|y|=1} ({}^ty\nabla^2\phi(x)y) \right],$$

with

$$\mathcal{A} = \left\{ \phi \in \mathcal{C}_b^2(\mathbb{R}^d) \mid G = \{\phi > 0\}, \quad \partial G = \{\phi = 0\}, \quad |\nabla\phi(x)| = 1, \quad \forall x \in \partial G \right\}.$$

Then the assumption $\alpha d < 2$ becomes $\alpha^* d < 2$. Moreover, α^* has some nice properties given by the following proposition.

Proposition 2.36 ¹

$$\begin{aligned} \alpha^* &= \inf_{\phi \in \mathcal{A}} \left[\sup_{x \in \text{co}(\bar{G}) \setminus \bar{G}} \sup_{|y|=1} ({}^ty\nabla^2\phi(x)y) \right], \\ \alpha^* &\geq \underline{\alpha}^* := \inf_{\phi \in \mathcal{A}} \left[\sup_{x \in \partial G} \sup_{|y|=1} ({}^ty\nabla^2\phi(x)y) \right], \\ \underline{\alpha}^* &= \sup_{x \in \partial G} \sup_{\substack{|y|=1 \\ y \in \{\nabla\phi(x)\}^\perp}} ({}^ty\nabla^2\phi(x)y), \quad \forall \phi \in \mathcal{A}. \end{aligned}$$

The proof is given in the Appendix.

Remark 2.37

- For some \bar{G} we can have $\alpha^* = \underline{\alpha}^*$, that is to say α^* only depends on the local geometry of ∂G characterised by $\underline{\alpha}^*$. For example, it is easy to show that in dimension 2, the domain given by figure 2.1 verifies such an equality.
- It is also possible to find a domain \bar{G} such that $\alpha^* > \underline{\alpha}^*$: in this case α^* depends also on the global geometry of \bar{G} . Indeed, let us consider the domain given by figure 2.2. This domain is a gathering of four arcs of a circle, so it is \mathcal{C}^1 but only \mathcal{C}^2 by pieces. This is not an important restriction because it is possible to mollify this domain near the four points where the boundary is not \mathcal{C}^2 . It is easy to see

¹This proposition and the remark that follows are not in the article [77]. The author would like to thanks Michel Pierre for an helpful conversation that led him to these results.

that $\underline{\alpha}^*$ does not depend on $\nu \in]\nu_0, 2\pi[$ where ν_0 is the angle such that $A = B$. Moreover, for any $\phi \in \mathcal{A}$, we can consider the function

$$\psi : \begin{array}{ll} [0, 1] & \rightarrow \mathbb{R}^- \\ t & \mapsto \phi(tx_A + (1-t)x_B). \end{array}$$

Since $\psi'(0) = -|x_B - x_A|$ and $\psi'(1) = |x_B - x_A|$, there exists $t \in]0, 1[$ such that $\psi''(t) = 2|x_B - x_A|$. So there is a point $C \in [A, B]$ such that ${}^t y \nabla^2 \phi(x_C) y = \frac{2}{|x_B - x_A|}$ where $y = \frac{x_B - x_A}{|x_B - x_A|}$. Thus, $\alpha^* \rightarrow +\infty$ when $\nu \rightarrow \nu_0$.

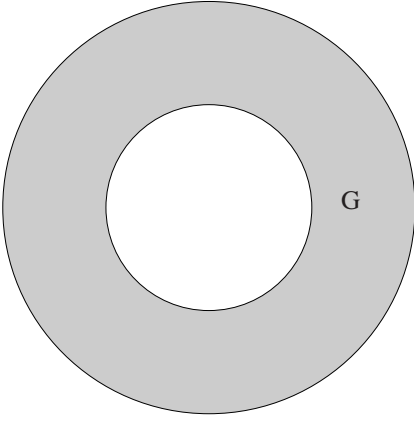


Figure 2.1: The “donut” example.

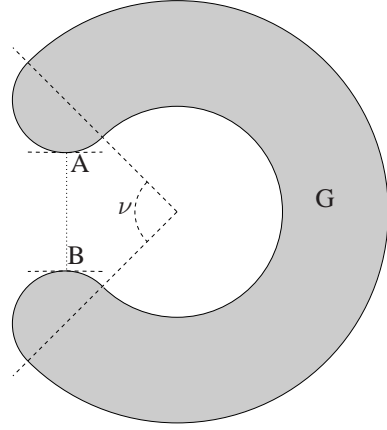


Figure 2.2: The “eaten donut” example.

2.8 Appendix

2.8.1 Proof of Proposition 2.18.

We will prove that for all Lipschitz functions f , $\mathbb{E}_{\nu_n}[f] \xrightarrow{n \rightarrow +\infty} \mathbb{E}_\nu[f]$. We have, for all $t \geq 0$,

$$\begin{aligned} |\mathbb{E}_{\nu_n}[f] - \mathbb{E}_\nu[f]| &= \left| \int_{\overline{G}} \mathbb{E}[f(X_t^{n,x})] \nu_n(dx) - \int_{\overline{G}} \mathbb{E}[f(X_t^x)] \nu(dx) \right| \\ &\leq \underbrace{\left| \int_{\overline{G}} \mathbb{E}[f(X_t^{n,x})] \nu_n(dx) - \int_{\overline{G}} \mathbb{E}[f(X_t^{n,y})] \nu(dy) \right|}_{A_{n,t}} \\ &\quad + \underbrace{\left| \int_{\overline{G}} \mathbb{E}[f(X_t^{n,x}) - f(X_t^x)] \nu(dx) \right|}_{B_{n,t}}. \end{aligned}$$

Firstly,

$$\begin{aligned} A_{n,t} &\leq \int_{\overline{G}} \int_{\overline{G}} \mathbb{E}[|f(X_t^{n,x}) - f(X_t^{n,y})|] \nu_n(dx) \nu(dy) \\ &\leq K_f \int_{\overline{G}} \int_{\overline{G}} \mathbb{E}[|X_t^{n,x} - X_t^{n,y}|] \nu_n(dx) \nu(dy). \end{aligned}$$

$\nabla^2 U_n \geq cI$, so $-\nabla U_n$ is dissipative : we can prove that (see e.g. Proposition 3.3 in [43])

$$\mathbb{E} |X_t^{n,x} - X_t^{n,y}| \leq e^{-ct} |x - y|.$$

Then, by simple computations

$$\begin{aligned} \int_{\overline{G}} \int_{\overline{G}} |x - y| \nu_n(dx) \nu(dy) &\leq \int_{\overline{G}} |x| \nu_n(dx) + \int_{\overline{G}} |y| \nu(dy) \\ &\leq \frac{1}{N} \int_{\mathbb{R}^d} |x| e^{-U(x)} dx + \int_{\overline{G}} |y| \nu(dy) < +\infty, \end{aligned}$$

with N defined in Corollary 2.19. So, $A_{n,t} \leq C e^{-ct} \xrightarrow{t \rightarrow +\infty} 0$, and the limit is uniform in n . Moreover,

$$B_{n,t} \leq K_f \int_{\overline{G}} \mathbb{E} |X_t^{n,x} - X_t^x| \nu(dx) \leq K_f \int_{\overline{G}} \mathbb{E} \left[\sup_{0 \leq s \leq t} |X_s^{n,x} - X_s^x| \right] \nu(dx).$$

So, by Theorem 1 in [69], $B_{n,t} \xrightarrow{n \rightarrow +\infty} 0$ when t is fixed. In conclusion, for all $t > 0$,

$$\limsup_{n \rightarrow +\infty} |\mathbb{E}_{\nu_n}[f] - \mathbb{E}_{\nu}[f]| \leq C e^{-ct}.$$

So we can conclude the proof by letting $t \rightarrow +\infty$. □

2.8.2 Proof of Corollary 2.19.

For all $f \in \mathcal{C}_{lip}^0(\mathbb{R}^d)$ we have

$$\mathbb{E}_{\nu_n}[f] = \frac{1}{N_n} \int_{\mathcal{R}^d} f(x) e^{-U_n(x)} dx.$$

According to Proposition 2.18, $\mathbb{E}_{\nu_n}[f] \xrightarrow{n \rightarrow +\infty} \mathbb{E}_{\nu}[f]$, and by the dominated convergence theorem, we show that

$$\frac{1}{N_n} \int_{\mathcal{R}^d} f(x) e^{-U_n(x)} dx \xrightarrow{n \rightarrow +\infty} \frac{1}{N} \int_{\overline{G}} f(x) e^{-U(x)} dx.$$

So

$$\mathbb{E}_{\nu}[f] = \frac{1}{N} \int_{\overline{G}} f(x) e^{-U(x)} dx.$$

Then the first claim ends by the use of a density argument and the monotone class theorem. Finally, we are able to apply the dominated convergence theorem to show that ν_n converges narrowly to ν . □

2.8.3 Proof of Proposition 2.21.

We know that $\nabla^2 U_n \geq cI$. So, according to the Bakry-Emery criterion (see [5]), we have the Poincaré inequality

$$\text{Var}_{\nu_n}(f) \leq -c^{-1} \langle \mathcal{L}_n f, f \rangle, \quad \forall f \in \mathbb{D}_2(\mathcal{L}_n).$$

Now, we are allowed to use Theorem 3.1 in [52]:

$$\mathbb{P}_{\nu} \left(-\frac{1}{T} \int_0^T \mathcal{L} \phi(X_s^n) ds \leq -\mathbb{E}_{\nu_n}[\mathcal{L} \phi] - \varepsilon \right) \leq \mathbb{E}_{\nu} \left[\left(\frac{d\nu}{d\nu_n} \right)^2 \right]^{1/2} \exp \left(-\frac{c\varepsilon^2 T}{\delta^2} \right).$$

Firstly, by dominated convergence theorem

$$\mathbb{E}_{\nu} \left[\left(\frac{d\nu}{d\nu_n} \right)^2 \right]^{1/2} = \frac{N_n}{N} \xrightarrow{n \rightarrow +\infty} 1.$$

Moreover, applying Proposition 2.18,

$$\mathbb{E}_{\nu_n}[\mathcal{L} \phi] \xrightarrow{n \rightarrow +\infty} \mathbb{E}_{\nu}[\mathcal{L} \phi].$$

Finally,

$$\int_{\overline{G}} \mathbb{E} \left| \frac{1}{T} \int_0^T \mathcal{L} \phi(X_s^{n,x}) ds - \frac{1}{T} \int_0^T \mathcal{L} \phi(X_s^x) ds \right| \nu(dx) \leq K_{\mathcal{L} \phi} \int_{\overline{G}} \mathbb{E} \left[\sup_{s \in [0, T]} |X_s^{n,x} - X_s^x| \right] \nu(dx).$$

But, according to [69],

$$\mathbb{E} \left[\sup_{s \in [0, T]} |X_s^{n, x} - X_s^x| \right] \xrightarrow{n \rightarrow +\infty} 0$$

and the limit is uniform in x belonging to \bar{G} . So

$$\int_{\bar{G}} \mathbb{E} \left| \frac{1}{T} \int_0^T \mathcal{L}\phi(X_s^{n, x}) ds - \frac{1}{T} \int_0^T \mathcal{L}\phi(X_s^x) ds \right| \nu(dx) \xrightarrow{n \rightarrow +\infty} 0,$$

and, as convergence in L^1 implies convergence in law, the claim follows. \square

2.8.4 Proof of Proposition 2.36.

First of all, it is easy to see that

$$\alpha^* \geq \inf_{\phi \in \mathcal{A}} \left[\sup_{x \in co(\bar{G}) \setminus \bar{G}} \sup_{|y|=1} ({}^t y \nabla^2 \phi(x) y) \right].$$

To show the equality, we will do a proof by contradiction. Let us assume that there exists $\tilde{\phi}$ such that

$$\sup_{x \in co(\bar{G}) \setminus \bar{G}} \sup_{|y|=1} ({}^t y \nabla^2 \tilde{\phi}(x) y) < \inf_{\phi \in \mathcal{A}} \left[\sup_{x \in co(\bar{G})} \sup_{|y|=1} ({}^t y \nabla^2 \phi(x) y) \right],$$

and let us consider $h_\varepsilon : \mathbb{R} \rightarrow \mathbb{R}$ a \mathcal{C}^2 function such that

- $h_\varepsilon(x) = x$ for $x \leq \varepsilon$,
- $h_\varepsilon(x) = 2\varepsilon$ for $x \geq 2\varepsilon$,
- $0 \leq h'_\varepsilon(x) \leq 1$ and $h''_\varepsilon(x) \leq 0$ for $x \in [\varepsilon, 2\varepsilon]$.

We set $\phi_\varepsilon = h_\varepsilon \circ \tilde{\phi}$. Then

$$\inf_{\phi \in \mathcal{A}} \left[\sup_{x \in co(\bar{G})} \sup_{|y|=1} ({}^t y \nabla^2 \phi(x) y) \right] \leq \sup_{x \in co(\bar{G})} \sup_{|y|=1} ({}^t y \nabla^2 \phi_\varepsilon(x) y),$$

and

$$\sup_{x \in co(\bar{G})} \sup_{|y|=1} ({}^t y \nabla^2 \phi_\varepsilon(x) y) = \sup_{x \in E_{2\varepsilon}} \sup_{|y|=1} ({}^t y \nabla^2 \phi_\varepsilon(x) y),$$

with $E_{2\varepsilon} = co(\bar{G}) \cap \tilde{\phi}^{-1}(] - \infty, 2\varepsilon])$. But, we have

$$\begin{aligned} 0 &< \sup_{x \in E_{2\varepsilon}} \sup_{|y|=1} ({}^t y \nabla^2 \phi_\varepsilon(x) y) &= \sup_{|y|=1} ({}^t y \nabla^2 \phi_\varepsilon(x_\varepsilon) y) \\ &= \underbrace{h''_\varepsilon(\tilde{\phi}(x_\varepsilon)) \sup_{|y|=1} ({}^t y \nabla^2 \tilde{\phi}(x_\varepsilon) {}^t \nabla \tilde{\phi}(x_\varepsilon) y)}_{\leq 0} \\ &\quad + \underbrace{h'_\varepsilon(\tilde{\phi}(x_\varepsilon)) \sup_{|y|=1} ({}^t y \nabla^2 \tilde{\phi}(x_\varepsilon) y)}_{\substack{\leq 1 \\ \geq 0}} \\ &\leq \sup_{|y|=1} ({}^t y \nabla^2 \tilde{\phi}(x_\varepsilon) y) \\ &\leq \sup_{x \in E_{2\varepsilon}} \sup_{|y|=1} ({}^t y \nabla^2 \tilde{\phi}(x) y). \end{aligned}$$

Moreover, $\nabla^2 \tilde{\phi}$ is a continuous function, so

$$\sup_{x \in E_{2\varepsilon}} \sup_{|y|=1} ({}^t y \nabla^2 \tilde{\phi}(x) y) \xrightarrow{\varepsilon \rightarrow 0} \sup_{x \in co(\bar{G}) \setminus \bar{G}} \sup_{|y|=1} ({}^t y \nabla^2 \tilde{\phi}(x) y),$$

and then

$$\inf_{\phi \in \mathcal{A}} \left[\sup_{x \in co(\bar{G})} \sup_{|y|=1} ({}^t y \nabla^2 \phi(x) y) \right] \leq \sup_{x \in co(\bar{G}) \setminus \bar{G}} \sup_{|y|=1} ({}^t y \nabla^2 \tilde{\phi}(x) y),$$

which is a contradiction: we obtain the first equality.

Now, let us show the end of the proposition. Obviously we have

$$\alpha^* \geq \underline{\alpha}^* \geq \inf_{\phi \in \mathcal{A}} \sup_{x \in \partial G} \sup_{\substack{|y|=1 \\ y \in \{\nabla \phi(x)\}^\perp}} ({}^t y \nabla^2 \phi(x) y).$$

Moreover, the quantity

$$\sup_{x \in \partial G} \sup_{\substack{|y|=1 \\ y \in \{\nabla \phi(x)\}^\perp}} ({}^t y \nabla^2 \phi(x) y)$$

does not depends on $\phi \in \mathcal{A}$, so

$$\underline{\alpha}^* \geq \sup_{x \in \partial G} \sup_{\substack{|y|=1 \\ y \in \{\nabla \phi(x)\}^\perp}} ({}^t y \nabla^2 \phi(x) y), \quad \forall \phi \in \mathcal{A}.$$

To show the equality, we will take the oriented distance function $\phi = d(\cdot, \bar{G}^c) - d(\cdot, \bar{G})$. Thanks to Theorem 5.6 in [37], we know that there exists an open neighborhood of ∂G where ϕ is \mathcal{C}^2 . Outside of this neighborhood, it is possible to consider a mollification of ϕ to obtain a \mathcal{C}^2 function on \mathbb{R}^d . Moreover, it is proved in section 5.5 of the article [37] that

$$\nabla^2 \phi(x) \nabla \phi(x) = 0, \quad \forall x \in \partial G.$$

So, in an orthogonal basis such that $\nabla \phi(x)$ is the first vector, we have

$$\nabla^2 \phi(x) = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & M(x) & & \\ 0 & & & \end{bmatrix}.$$

Then

$$\sup_{x \in \partial G} \sup_{|y|=1} ({}^t y \nabla^2 \phi(x) y) = \sup_{x \in \partial G} \sup_{\substack{|y|=1 \\ y \in \{\nabla \phi(x)\}^\perp}} ({}^t y \nabla^2 \phi(x) y).$$

□

Acknowledgements. The author would like to thank his Ph.D. advisers, Philippe Briand and Ying Hu, an anonymous referee and the associate editor for their careful reading and helpful comments.

Chapitre 3

Quelques résultats complémentaires sur les EDSREs

Dans ce chapitre nous ajoutons quelques résultats concernant les EDSREs ne se trouvant pas dans l'article [77].

3.1 Un résultat d'existence et d'unicité concernant les EDSRs généralisées en horizon infini

3.1.1 Motivation

Revenons au problème initial que nous cherchons à résoudre dans le chapitre 2. Nous souhaitons montrer l'existence d'une solution (Y, Z, μ) pour l'EDSRE

$$Y_t^x = Y_T^x + \int_t^T \psi(X_s^x, Z_s^x) ds + \int_t^T [g(X_s^x) - \mu] dK_s^x - \int_t^T Z_s^x dW_s.$$

Pour cela, l'idée la plus naturelle est de traduire en termes probabilistes la démonstration de l'article [11] qui consiste à étudier la limite, lorsque $\alpha \rightarrow 0$, de la solution v^α de l'EDP

$$\begin{cases} \mathcal{L}v^\alpha(x) + \psi(x, {}^t \nabla v^\alpha(x) \sigma(x)) - \alpha^2 v^\alpha(x) = 0, & x \in G \\ \frac{\partial v^\alpha}{\partial n}(x) + g(x) - \alpha v^\alpha(x) = 0, & x \in \partial G, \end{cases}$$

avec

$$\mathcal{L}f(x) = \frac{1}{2} \text{Tr}(\sigma(x) {}^t \sigma(x) \nabla^2 f(x)) + {}^t b(x) \nabla f(x).$$

D'un point de vue probabiliste, cela revient à introduire l'EDSR

$$Y_t^{x,\alpha} = Y_T^{x,\alpha} + \int_t^T [\psi(X_s^x, Z_s^{x,\alpha}) - \alpha^2 Y_s^{x,\alpha}] ds + \int_t^T [g(X_s^x) - \alpha Y_s^{x,\alpha}] dK_s^x - \int_t^T Z_s^{x,\alpha} dW_s. \quad (3.1)$$

En premier lieu, il convient de savoir si cette EDSR possède une solution éventuellement unique. Ce type d'équation a déjà été envisagé, pour Y en dimension quelconque, dans l'article [75], mais l'étude proposée fait apparaître des hypothèses techniques dont nous souhaitons nous affranchir. En utilisant les spécificités de la dimension 1, Briand et Hu [23] proposent un résultat qui va dans ce sens mais ces derniers se cantonnent aux EDSRs non généralisées. Il est néanmoins possible de reprendre leur démonstration pour l'adapter à notre situation.

3.1.2 Le résultat d'existence et d'unicité

Considérons l'EDSR généralisée en horizon infini

$$Y_t = Y_T + \int_t^T f(s, Y_s, Z_s) ds + \int_t^T g(s, Y_s) dA_s - \int_t^T Z_s dW_s, \quad \forall 0 \leq t \leq T. \quad (3.2)$$

f est une fonction aléatoire définie sur $[0, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^{1 \times d}$, à valeur dans \mathbb{R} et mesurable par rapport aux tribus $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}^{1 \times d})$ et $\mathcal{B}(\mathbb{R})$, \mathcal{P} désignant la tribu des événements prévisibles. g est une fonction aléatoire définie sur $[0, T] \times \Omega \times \mathbb{R}$, à valeur dans \mathbb{R} et mesurable par rapport aux tribus $\mathcal{P} \otimes \mathcal{B}(\mathbb{R})$ et $\mathcal{B}(\mathbb{R})$. De plus, $(A_s)_{s \geq 0}$ est un processus continu croissant, \mathcal{F}_t -progressivement mesurable, à valeur dans \mathbb{R} et vérifiant $A_0 = 0$. Enfin, pour tous $\alpha \in \mathbb{R}$ et $\beta \in \mathbb{R}$, on note $\mathcal{M}^{2, \alpha, \beta}(\mathbb{R}^+, E)$ l'espace de Hilbert des processus \mathcal{F}_t -progressivement mesurables à valeur dans l'espace euclidien E tels que

$$\|X\|_{\alpha, \beta}^2 := \mathbb{E} \left[\int_0^{+\infty} e^{\alpha s + \beta A_s} |X_s|^2 ds \right] < \infty.$$

Considérons les hypothèses suivantes sur les fonctions f et g :

(H). Il existe quatre constantes $M_f \geq 0$, $M_g \geq 0$, $\mu > 0$ et $\mu' > 0$ telles que, $d\mathbb{P} \otimes dt$ -p.s.,

1. f est M_f -lipschitzienne en y et z : pour tous $(y_1, z_1), (y_2, z_2)$ dans $\mathbb{R} \times \mathbb{R}^d$,
 $|f(t, y_1, z_1) - f(t, y_2, z_2)| \leq M_f(|y_1 - y_2| + |z_1 - z_2|)$;
2. g est M_g -lipschitzienne en y : pour tous y_1, y_2 dans \mathbb{R} ,
 $|g(t, y_1) - g(t, y_2)| \leq M_g|y_1 - y_2|$;
3. f et g sont monotones en y : pour tous y_1, y_2 dans \mathbb{R} et z dans \mathbb{R}^d on a
 $(y_1 - y_2) \cdot (f(t, y_1, z) - f(t, y_2, z)) \leq -\mu(y_1 - y_2)^2$ et
 $(y_1 - y_2) \cdot (g(t, y_1) - g(t, y_2)) \leq -\mu'(y_1 - y_2)^2$;
4. $|f(t, 0, 0)| \leq M_f$ et $|g(t, 0)| \leq M_g$.

Théorème 3.1 *Supposons que les hypothèses (H) sont vérifiées. Alors l'EDSR en horizon infini généralisée (3.2) possède une solution (Y, Z) qui appartient à $\mathcal{M}^{2, -2\mu, -2\mu'}(\mathbb{R}^+, \mathbb{R} \times \mathbb{R}^d)$ pour tout $\varepsilon > 0$ et telle que Y est borné. Cette solution est unique parmi les processus (Y, Z) tels que Y est continu borné et Z appartient à $\mathcal{M}_{loc}^2(\mathbb{R}^+, \mathbb{R}^{1 \times d})$. De plus on a l'estimation suivante :*

$$|Y| \leq \left(\frac{M_f}{\mu} \vee \frac{M_g}{\mu'} \right).$$

Démonstration. Prouvons tout d'abord l'unicité. Supposons que (Y^1, Z^1) et (Y^2, Z^2) sont deux solutions de l'EDSR (3.2) telles que Y^1 et Y^2 sont bornés et continus et que Z^1 et Z^2 sont dans $\mathcal{M}_{loc}^2(\mathbb{R}^+, \mathbb{R}^{1 \times d})$. Posons $\tilde{Y} := Y^1 - Y^2$ et $\tilde{Z} := Z^1 - Z^2$. Ainsi \tilde{Y} est un processus continu et il existe $M > 0$ tel que $|\tilde{Y}|$ est \mathbb{P} -p.s. borné par M . On a alors

$$\tilde{Y}_t = \tilde{Y}_T + \int_t^T [f(s, Y_s^1, Z_s^1) - f(s, Y_s^2, Z_s^2)] ds + \int_t^T [g(s, Y_s^1) - g(s, Y_s^2)] dA_s - \tilde{Z}_s dW_s.$$

Afin de linéariser cette EDSR, nous définissons trois nouveaux processus (α, β, γ) , à valeur dans $\mathbb{R} \times \mathbb{R}^d \times \mathbb{R}$, en posant

$$\begin{aligned} \alpha_t &= \begin{cases} \frac{f(t, Y_t^1, Z_t^1) - f(t, Y_t^2, Z_t^1)}{Y_t^1 - Y_t^2} & \text{si } Y_t^1 - Y_t^2 \neq 0, \\ -\mu & \text{sinon,} \end{cases} \\ \beta_t &= \begin{cases} \frac{f(t, Y_t^2, Z_t^1) - f(t, Y_t^2, Z_t^2)}{|\tilde{Z}_t|^2} \tilde{Z}_t & \text{si } \tilde{Z}_t \neq 0, \\ 0 & \text{sinon,} \end{cases} \\ \gamma_t &= \begin{cases} \frac{g(t, Y_t^1) - g(t, Y_t^2)}{Y_t^1 - Y_t^2} & \text{si } Y_t^1 - Y_t^2 \neq 0, \\ -\mu' & \text{sinon.} \end{cases} \end{aligned}$$

Remarquons que (α, β, γ) sont trois processus bornés car f et g sont lipschitziennes. Finalement, on obtient

$$\tilde{Y}_t = \tilde{Y}_T + \int_t^T \alpha_s \tilde{Y}_s ds + \int_t^T \gamma_s \tilde{Y}_s dA_s - \int_t^T \tilde{Z}_s (dW_s - \beta_s ds).$$

Posons, à t fixé, $R_s = \exp(\int_t^s \alpha_u du + \int_t^s \gamma_u dA_u)$, et appliquons la formule d'Itô au processus $R_s \tilde{Y}_s$:

$$R_n \tilde{Y}_n - R_t \tilde{Y}_t = \int_t^n R_s \tilde{Z}_s (dW_s - \beta_s ds),$$

ce qui se réécrit,

$$\tilde{Y}_t = R_n \tilde{Y}_n - \int_t^n R_s \tilde{Z}_s d\tilde{W}_s, \quad (3.3)$$

avec $\tilde{W}_t = W_t - \int_0^t \beta_s ds$. Soit \mathbb{Q}_n la mesure de probabilité sur (Ω, \mathcal{F}_n) dont la densité par rapport à $\mathbb{P}|_{\mathcal{F}_n}$ est

$$\exp \left(\int_0^n \beta_s dW_s - \frac{1}{2} \int_0^n |\beta_s|^2 ds \right).$$

β est un processus borné donc la condition de Novikov est vérifiée et l'on peut appliquer le théorème de Girsanov : Les mesures de probabilité \mathbb{Q}_n et $\mathbb{P}|_{\mathcal{F}_n}$ sont mutuellement absolument continues et $(\tilde{W}_t)_{0 \leq t \leq n}$ est un mouvement brownien sous \mathbb{Q}_n . En prenant l'espérance conditionnelle par rapport à \mathcal{F}_t de l'équation (3.3), nous obtenons

$$|\tilde{Y}_t| \leq \mathbb{E}^{\mathbb{Q}_n} [R_n |\tilde{Y}_n| | \mathcal{F}_t], \quad \mathbb{Q}_n\text{-p.s.} \quad (3.4)$$

Or l'hypothèse de monotonie portant sur f et g nous donne $\alpha \leq -\mu$ et $\gamma \leq -\mu' < 0$ $d\mathbb{P} \otimes dt$ -p.s.. Donc $R_n \leq e^{-\mu(n-t) - \mu'(A_n - A_t)} \leq e^{-\mu(n-t)}$ $\mathbb{P}|_{\mathcal{F}_n}$ -p.s.. De plus, $|\tilde{Y}|$ est $\mathbb{P}|_{\mathcal{F}_n}$ -p.s. borné par $2M$ et $\mathbb{P}|_{\mathcal{F}_n} \sim \mathbb{Q}_n$, donc

$$R_n |\tilde{Y}_n| \leq 2M e^{-\mu(n-t)} \quad \mathbb{Q}_n\text{-p.s.}$$

En revenant à (3.4), nous obtenons

$$|\tilde{Y}_t| \leq 2M e^{-\mu(n-t)} \quad \mathbb{Q}_n\text{-p.s.}$$

Comme $\mathbb{P}|_{\mathcal{F}_n} \sim \mathbb{Q}_n$ nous pouvons en déduire que

$$\mathbb{P}\text{-p.s.} \quad \forall n \in \mathbb{N}, \quad n \geq t, \quad |\tilde{Y}_t| \leq 2M e^{-\mu(n-t)}.$$

Il reste à faire tendre n vers l'infini pour obtenir

$$|\tilde{Y}_t| = 0 \quad \mathbb{P}\text{-p.s.}$$

Enfin, la continuité des processus Y^1 et Y^2 nous assure que $Y^1 = Y^2$ \mathbb{P} -p.s.. Pour terminer la preuve de l'unicité, nous appliquons la formule d'Itô au processus $|\tilde{Y}_t|^2$:

$$\begin{aligned} |\tilde{Y}_t|^2 - |\tilde{Y}_0|^2 &= -2 \int_0^t [f(s, Y_s^1, Z_s^1) - f(s, Y_s^2, Z_s^2)] \tilde{Y}_s ds - 2 \int_0^t [g(s, Y_s^1) - g(s, Y_s^2)] \tilde{Y}_s dA_s \\ &\quad + \int_0^t |\tilde{Z}_s|^2 ds + \int_0^t \tilde{Z}_s \tilde{Y}_s dW_s, \end{aligned}$$

ce qui nous donne finalement,

$$\mathbb{E} \left[\int_0^t |\tilde{Z}_s|^2 ds \right] = 0, \quad \forall t \geq 0.$$

Passons maintenant à la démonstration de l'existence. On note (Y^n, Z^n) l'unique solution de l'EDSR généralisée

$$Y_t^n = \int_t^n f(s, Y_s^n, Z_s^n) ds + \int_t^n g(s, Y_s^n) dA_s - \int_t^n Z_s^n dW_s, \quad (3.5)$$

vérifiant

$$\mathbb{E} \left[\sup_{0 \leq t \leq n} |Y_t^n|^2 + \int_0^n |Z_t^n|^2 dt \right] < +\infty.$$

L'existence et l'unicité de cette solution sont établies dans l'article [75]. Dans un premier temps nous allons obtenir une estimation a priori pour Y^n . Comme pour la preuve de l'unicité, nous posons

$$\alpha_t^n = \begin{cases} \frac{f(t, Y_t^n, Z_t^n) - f(t, 0, Z_t^n)}{Y_t^n} & \text{si } Y_t^n \neq 0, \\ -\mu & \text{sinon,} \end{cases}$$

$$\beta_t^n = \begin{cases} \frac{f(t, 0, Z_t^n) - f(t, 0, 0)}{|Z_t^n|^2} Z_t^n & \text{si } Z_t^n \neq 0, \\ 0 & \text{sinon,} \end{cases}$$

$$\gamma_t^n = \begin{cases} \frac{g(t, Y_t^n) - g(t, 0)}{Y_t^n} & \text{si } Y_t^n \neq 0, \\ -\mu' & \text{sinon.} \end{cases}$$

L'EDSR linéarisée est alors donnée par

$$Y_t^n = \int_t^n [\alpha_s^n Y_s^n + Z_s^n \beta_s^n + f(s, 0, 0)] ds + \int_t^n [\gamma_s^n Y_s^n + g(s, 0)] dA_s - \int_t^n Z_s^n dW_s. \quad (3.6)$$

On définit

$$R_u^n = \exp \left(\int_t^u \alpha_s^n ds + \int_t^u \gamma_s^n dA_s \right) \text{ et } \tilde{W}_t^n = W_t - \int_0^t \beta_s^n ds.$$

En appliquant la formule d'Itô au processus $R_n(s)Y_n(s)$ nous obtenons

$$Y_t^n = \int_t^n R_s^n f(s, 0, 0) ds + \int_t^n R_s^n g(s, 0) dA_s - \int_t^n R_s^n Z_s^n d\tilde{W}_s^n.$$

On introduit la mesure de probabilité \mathbb{Q}_n sur (Ω, \mathcal{F}_n) dont la densité par rapport à $\mathbb{P}|_{\mathcal{F}_n}$ est donnée par

$$\exp \left(\int_0^n \beta_s^n dW_s - \frac{1}{2} \int_0^n |\beta_s^n|^2 ds \right).$$

β^n est un processus borné donc la condition de Novikov est vérifiée et nous pouvons appliquer le théorème de Girsanov : Les mesures de probabilité \mathbb{Q}_n et $\mathbb{P}|_{\mathcal{F}_n}$ sont mutuellement absolument continues et $(\tilde{W}_t^n)_{0 \leq t \leq n}$ est un mouvement brownien sous \mathbb{Q}_n . Nous en déduisons

$$|Y_t^n| \leq \mathbb{E}^{\mathbb{Q}_n} \left[\int_t^n R_s^n |f(s, 0, 0)| ds + \int_t^n R_s^n |g(s, 0)| dA_s \middle| \mathcal{F}_t \right], \quad \mathbb{Q}_n\text{-p.s.}$$

En remarquant que $|f(s, 0, 0)| \leq M_f$ et que $R_s^n \leq e^{-\mu(s-t) - \mu'(A_s - A_t)}$, nous avons

$$\begin{aligned} |Y_t^n| &\leq \mathbb{E}^{\mathbb{Q}_n} \left[M_f \int_t^n e^{-\mu(s-t) - \mu'(A_s - A_t)} ds + M_g \int_t^n e^{-\mu(s-t) - \mu'(A_s - A_t)} dA_s \middle| \mathcal{F}_t \right] \\ &\leq \left(\frac{M_f}{\mu} \vee \frac{M_g}{\mu'} \right) \mathbb{E}^{\mathbb{Q}_n} \left[\int_t^n (\mu e^{-\mu(s-t) - \mu'(A_s - A_t)} ds + \mu' e^{-\mu(s-t) - \mu'(A_s - A_t)} dA_s) \middle| \mathcal{F}_t \right] \\ &\leq \left(\frac{M_f}{\mu} \vee \frac{M_g}{\mu'} \right), \quad \mathbb{Q}_n\text{-p.s.} \end{aligned}$$

Comme Y^n est un processus continu et que $\mathbb{P}|_{\mathcal{F}_n} \sim \mathbb{Q}_n$, nous obtenons finalement

$$\mathbb{P}\text{-p.s.} \quad \forall n \in \mathbb{N}, \quad \forall t \in [0, n], \quad |Y_t^n| \leq \left(\frac{M_f}{\mu} \vee \frac{M_g}{\mu'} \right). \quad (3.7)$$

Nous allons maintenant étudier la convergence de la suite $(Y^n, Z^n)_{n \in \mathbb{N}}$ dans l'espace de Hilbert

$$\mathcal{M}^{2, -2\mu, -2\varepsilon}(\mathbb{R}^+, \mathbb{R} \times \mathbb{R}^d)$$

pour $\varepsilon > 0$ fixé. On définit Y^n et Z^n sur \mathbb{R}^+ en posant $Y^n(t) = 0$ et $Z^n(t) = 0$ pour $t > n$. Fixons $t \leq n \leq m$ et posons $\tilde{Y} := Y^m - Y^n$ et $\tilde{Z} := Z^m - Z^n$. La formule d'Itô nous donne

$$\begin{aligned} \tilde{Y}_t &= \int_t^m [f(s, Y_s^m, Z_s^m) - \mathbb{1}_{s \leq n} f(s, Y_s^n, Z_s^n)] ds \\ &\quad + \int_t^m [g(s, Y_s^m) - \mathbb{1}_{s \leq n} g(s, Y_s^n)] dA_s - \int_t^m \tilde{Z}_s dW_s \end{aligned}$$

Comme pour la preuve de l'unicité et de l'estimation a priori, nous allons utiliser le même type de linéarisation en posant

$$\begin{aligned}\alpha_t^{n,m} &= \begin{cases} \frac{f(t, Y_t^m, Z_t^m) - f(t, Y_t^n, Z_t^n)}{Y_t^m - Y_t^n} & \text{si } Y_t^m - Y_t^n \neq 0, \\ -\mu & \text{sinon,} \end{cases} \\ \beta_t^{n,m} &= \begin{cases} \frac{f(t, Y_t^n, Z_t^m) - f(t, Y_t^n, Z_t^n)}{|\tilde{Z}_t|^2} \tilde{Z}_t & \text{si } \tilde{Z}_t \neq 0, \\ 0 & \text{sinon,} \end{cases} \\ \gamma_t^{m,n} &= \begin{cases} \frac{g(t, Y_t^m) - g(t, Y_t^n)}{Y_t^m - Y_t^n} & \text{si } Y_t^m - Y_t^n \neq 0, \\ -\mu' & \text{sinon.} \end{cases}\end{aligned}$$

On a alors

$$\begin{aligned}\tilde{Y}_t &= \int_t^m \left(\alpha_s^{n,m} \tilde{Y}_s + \tilde{Z}_s \beta_s^{n,m} + \mathbb{1}_{s>n} f(s, 0, 0) \right) ds \\ &\quad + \int_t^m \left(\gamma_s^{n,m} \tilde{Y}_s + \mathbb{1}_{s>n} g(s, 0) \right) dA_s - \int_t^m \tilde{Z}_s dW_s.\end{aligned}$$

On définit

$$R_u^{n,m} = \exp \left(\int_t^u \alpha_s^{n,m} ds + \int_t^u \gamma_s^{n,m} dA_s \right) \text{ et } \tilde{W}_t^{n,m} = W_t - \int_0^t \beta_s^{n,m} ds.$$

En appliquant la formule d'Itô au processus $R_s^{n,m} \tilde{Y}_s$ nous obtenons

$$\tilde{Y}_t = \int_n^m R_s^{n,m} f(s, 0, 0) ds + \int_n^m R_s^{n,m} g(s, 0) dA_s - \int_n^m R_s^{n,m} \tilde{Z}_s d\tilde{W}_s^{n,m}.$$

On introduit alors la mesure de probabilité $\mathbb{Q}_{n,m}$ sur (Ω, \mathcal{F}_m) dont la densité par rapport à $\mathbb{P}|_{\mathcal{F}_m}$ est donnée par

$$\exp \left(\int_0^m \beta_s^{n,m} dW_s - \frac{1}{2} \int_0^m |\beta_s^{n,m}|^2 ds \right).$$

En appliquant une nouvelle fois le théorème de Girsanov nous en déduisons

$$\begin{aligned}|\tilde{Y}_t| &\leq \mathbb{E}^{\mathbb{Q}_{n,m}} \left[M_f \int_t^n e^{-\mu(s-t) - \mu'(A_s - A_t)} ds + M_g \int_t^n e^{-\mu(s-t) - \mu'(A_s - A_t)} dA_s \middle| \mathcal{F}_t \right], \quad \mathbb{Q}_{n,m}\text{-p.s.} \\ &\leq \left(\frac{M_f}{\mu} \vee \frac{M_g}{\mu'} \right) \mathbb{E}^{\mathbb{Q}_{n,m}} \left[\int_n^m (\mu e^{-\mu(s-t) - \mu'(A_s - A_t)} ds \right. \\ &\quad \left. + \mu' e^{-\mu(s-t) - \mu'(A_s - A_t)} dA_s \right) \middle| \mathcal{F}_t \right], \quad \mathbb{Q}_{n,m}\text{-p.s.} \\ &\leq \left(\frac{M_f}{\mu} \vee \frac{M_g}{\mu'} \right) e^{-\mu(n-t)}, \quad \mathbb{Q}_{n,m}\text{-p.s.}\end{aligned}$$

Comme \tilde{Y} est un processus continu et que $\mathbb{P}|_{\mathcal{F}_m} \sim \mathbb{Q}_{n,m}$, nous obtenons finalement

$$\mathbb{P}\text{-p.s.} \quad \forall 0 \leq t \leq n \leq m, \quad |Y_t^m - Y_t^n| \leq \left(\frac{M_f}{\mu} \vee \frac{M_g}{\mu'} \right) e^{-\mu(n-t)}. \quad (3.8)$$

D'après cette inégalité, pour tout $t \geq 0$ la suite de variables aléatoires $(Y_t^n)_{n \in \mathbb{N}}$ est une suite de Cauchy dans L^∞ et donc converge vers une limite notée Y_t , uniformément en t sur tout compact. En faisant tendre m vers l'infini, l'inégalité précédente devient

$$\mathbb{P}\text{-p.s.} \quad \forall 0 \leq t \leq n, \quad |Y_t - Y_t^n| \leq \left(\frac{M_f}{\mu} \vee \frac{M_g}{\mu'} \right) e^{-\mu(n-t)}. \quad (3.9)$$

Montrons maintenant que $(Y^n)_{n \in \mathbb{N}}$ est une suite de Cauchy dans $\mathcal{M}^{2, -2\mu, -2\varepsilon}(\mathbb{R}^+, \mathbb{R})$. On a

$$\mathbb{E} \left[\int_0^m e^{-2\mu s - 2\varepsilon A_s} |\tilde{Y}_s|^2 ds \right] = \mathbb{E} \left[\int_0^n e^{-2\mu s - 2\varepsilon A_s} |\tilde{Y}_s|^2 ds \right] + \mathbb{E} \left[\int_n^m e^{-2\mu s - 2\varepsilon A_s} |\tilde{Y}_s|^2 ds \right].$$

On applique alors l'inégalité (3.8) et l'estimation a priori (3.7) pour obtenir

$$\mathbb{E} \left[\int_0^m e^{-2\mu s - 2\varepsilon A_s} |\tilde{Y}_s|^2 ds \right] \leq \mathbb{E} \left[\int_0^n \left(\frac{M_f}{\mu} \vee \frac{M_g}{\mu'} \right)^2 e^{-2\mu n} ds \right] + \mathbb{E} \left[\int_n^m \left(\frac{M_f}{\mu} \vee \frac{M_g}{\mu'} \right)^2 e^{-2\mu s} ds \right],$$

c'est-à-dire,

$$\mathbb{E} \left[\int_0^m e^{-2\mu s - 2\varepsilon A_s} |Y_s^m - Y_s^n|^2 ds \right] \leq \left(\frac{M_f}{\mu} \vee \frac{M_g}{\mu'} \right)^2 e^{-2\mu n} \left(n + \frac{1}{2\mu} \right), \quad (3.10)$$

ce qui nous prouve que $(Y^n)_{n \in \mathbb{N}}$ est une suite de Cauchy dans $\mathcal{M}^{2, -2\mu, -2\varepsilon}(\mathbb{R}^+, \mathbb{R})$. Montrons maintenant la même chose pour le processus $(Z^n)_{n \in \mathbb{N}}$. Pour cela, on applique la formule d'Itô à $e^{-2\mu t - 2\varepsilon A_t} |\tilde{Y}_t|^2$ pour obtenir

$$\begin{aligned} \mathbb{E} \left[|\tilde{Y}_0|^2 + \int_0^m e^{-2\mu s - 2\varepsilon A_s} |\tilde{Z}_s|^2 ds \right] = \\ \mathbb{E} \left[\int_0^m e^{-2\mu s - 2\varepsilon A_s} \left(2\mu |\tilde{Y}_s|^2 + 2\tilde{Y}_s [f(s, Y_s^m, Z_s^m) - f(s, Y_s^n, Z_s^n)] \right) ds \right. \\ \left. + \int_0^m e^{-2\mu s - 2\varepsilon A_s} \left(2\varepsilon |\tilde{Y}_s|^2 + 2\tilde{Y}_s [g(s, Y_s^m) - g(s, Y_s^n)] \right) dA_s \right. \\ \left. + \int_n^m e^{-2\mu s - 2\varepsilon A_s} Y_s^m f(s, 0, 0) ds + \int_n^m e^{-2\mu s - 2\varepsilon A_s} Y_s^m g(s, 0) dA_s \right]. \end{aligned}$$

Les fonctions g et f vérifient l'hypothèse (\mathcal{H}) donc

$$\begin{aligned} \mathbb{E} \left[\int_0^m e^{-2\mu s - 2\varepsilon A_s} |\tilde{Z}_s|^2 ds \right] \leq \mathbb{E} \left[\int_0^m e^{-2\mu s - 2\varepsilon A_s} \left(2\mu |\tilde{Y}_s|^2 + 2M_f |\tilde{Y}_s|^2 + 2M_f |\tilde{Y}_s| |\tilde{Z}_s| \right) ds \right. \\ \left. + \int_0^m e^{-2\mu s - 2\varepsilon A_s} \left(2\varepsilon |\tilde{Y}_s|^2 + 2M_g |\tilde{Y}_s|^2 \right) dA_s \right. \\ \left. + \int_n^m e^{-2\mu s - 2\varepsilon A_s} M_f |Y_s^m| ds + \int_n^m e^{-2\mu s - 2\varepsilon A_s} M_g |Y_s^m| dA_s \right]. \end{aligned}$$

On utilise le fait que $2M_f |\tilde{Y}_s| |\tilde{Z}_s| \leq 2M_f^2 |\tilde{Y}_s|^2 + \frac{1}{2} |\tilde{Z}_s|^2$ pour obtenir

$$\begin{aligned} \frac{1}{2} \mathbb{E} \left[\int_0^m e^{-2\mu s - 2\varepsilon A_s} |\tilde{Z}_s|^2 ds \right] \leq \\ \mathbb{E} \left[\int_0^m e^{-2\mu s - 2\varepsilon A_s} (2\mu + 2M_f + 2M_f^2) |\tilde{Y}_s|^2 ds \right. \\ \left. + \int_0^m e^{-2\mu s - 2\varepsilon A_s} (2\varepsilon + 2M_g) |\tilde{Y}_s|^2 dA_s \right. \\ \left. + \int_n^m e^{-2\mu s - 2\varepsilon A_s} M_f |Y_s^m| ds + \int_n^m e^{-2\mu s - 2\varepsilon A_s} M_g |Y_s^m| dA_s \right]. \end{aligned}$$

Pour faciliter les calculs, notons A , B et C les trois termes du membre de droite de l'inégalité précédente. En utilisant l'inégalité (3.10), on obtient

$$A \leq (2\mu + 2M_f + 2M_f^2) \left(\frac{M_f}{\mu} \vee \frac{M_g}{\mu'} \right)^2 e^{-2\mu n} \left(n + \frac{1}{2\mu} \right).$$

De plus, l'inégalité (3.8) et l'estimation a priori (3.7) nous donnent

$$\begin{aligned} B &\leq (2\varepsilon + 2M_g) \mathbb{E} \left[\int_0^n e^{-2\mu s - 2\varepsilon A_s} |\tilde{Y}_s|^2 dA_s + \int_n^m e^{-2\mu n - 2\varepsilon A_s} |\tilde{Y}_s|^2 dA_s \right] \\ &\leq (2\varepsilon + 2M_g) \left(\frac{M_f}{\mu} \vee \frac{M_g}{\mu'} \right) \mathbb{E} \left[\int_0^n e^{-2\mu s - 2\varepsilon A_s} e^{-2\mu(n-s)} dA_s + \int_n^m e^{-2\mu n - 2\varepsilon A_s} dA_s \right] \\ &\leq (2\varepsilon + 2M_g) \left(\frac{M_f}{\mu} \vee \frac{M_g}{\mu'} \right) e^{-2\mu n} \frac{1}{2\varepsilon}. \end{aligned}$$

Enfin, nous avons

$$\begin{aligned} C &\leq \left(\frac{M_f}{\mu} \vee \frac{M_g}{\mu'} \right) \left(\frac{M_f}{\mu} \vee \frac{M_g}{\varepsilon} \right) \mathbb{E} \left[\int_n^m \frac{e^{-2\mu s - 2\varepsilon A_s}}{2} (2\mu ds + 2\varepsilon dA_s) \right] \\ &\leq \left(\frac{M_f}{\mu} \vee \frac{M_g}{\mu'} \right) \left(\frac{M_f}{\mu} \vee \frac{M_g}{\varepsilon} \right) \frac{e^{-2\mu n}}{2}, \end{aligned}$$

ce qui nous prouve que $(Y^n, Z^n)_{n \in \mathbb{N}}$ est une suite de Cauchy dans l'espace de Hilbert $\mathcal{M}^{2, -2\mu, -2\varepsilon}(\mathbb{R}^+, \mathbb{R} \times \mathbb{R}^d)$ et donc converge vers une limite (Y, Z) . Il ne reste plus qu'à prouver que ce couple de processus vérifie bien l'EDSR (3.2). Tout d'abord, l'estimation a priori (3.7) nous assure que

$$\mathbb{P}\text{-p.s.}, \quad \forall t \in \mathbb{R}^+, \quad |Y_t| \leq \left(\frac{M_f}{\mu} \vee \frac{M_g}{\mu'} \right).$$

Ensuite, pour $0 \leq r \leq t \leq n$, nous avons

$$Y_r^n - Y_t^n = \int_r^t f(s, Y_s^n, Z_s^n) ds + \int_r^t g(s, Y_s^n) dA_s - \int_r^t Z_s^n dW_s.$$

Ainsi, en appliquant le théorème de convergence dominé dans L^2 nous obtenons

$$Y_r - Y_t = \int_r^t f(s, Y_s, Z_s) ds + \int_r^t g(s, Y_s) dA_s - \int_r^t Z_s dW_s$$

ce qui termine la preuve. \square

3.1.3 Non application aux EDSREs

Comme nous le souhaitons, le théorème 3.1 permet d'obtenir l'existence d'une unique solution $(Y^{x, \alpha}, Z^{x, \alpha})$ pour l'EDSR généralisée en horizon infini (3.1). Il convient maintenant de s'intéresser aux processus limite lorsque α tend vers 0. La stratégie de démonstration présente dans les articles [11, 43] nécessite deux résultats distincts :

- le processus $|\alpha Y^{x, \alpha}|$ est borné uniformément en x et en α ,
- $x \mapsto Y_0^{x, \alpha}$ est une fonction lipschitz uniformément en α .

Considérons tout d'abord le premier résultat nécessaire. Si, comme dans l'article [43], nous souhaitons appliquer l'estimation a priori du théorème 3.1, alors nous obtenons que le processus $|\alpha^2 Y^{x, \alpha}|$ est borné ce qui n'est pas suffisant. Il est alors naturel de se demander s'il n'est pas possible de reprendre la démonstration du théorème 3.1 afin d'obtenir une estimation a priori plus forte. Dans ce but, regardons l'EDSR (3.2) lorsque $f(s, y, z) = 1 - \alpha^2 y$ et $g(s, y) = 1 - \alpha y$, la solution étant notée (Y^α, Z^α) . Une transformation classique permet d'obtenir de manière explicite la solution Y^α :

$$Y_t^\alpha = \mathbb{E} \left[\int_t^{+\infty} e^{-\alpha^2(s-t) - \alpha(A_s - A_t)} ds + \int_t^{+\infty} e^{-\alpha^2(s-t) - \alpha(A_s - A_t)} dA_s \middle| \mathcal{F}_t \right].$$

On peut tout d'abord remarquer que

$$\begin{aligned} \left| \mathbb{E} \left[\int_t^{+\infty} e^{-\alpha^2(s-t) - \alpha(A_s - A_t)} dA_s \middle| \mathcal{F}_t \right] \right| &\leq \mathbb{E} \left[\int_t^{+\infty} e^{-\alpha(A_s - A_t)} dA_s \middle| \mathcal{F}_t \right] \\ &\leq \frac{1}{\alpha}, \end{aligned}$$

ce qui est la bonne majoration. En ce qui concerne la première intégrale, étudions ce qui se passe sur des exemples. Si nous prenons $A_s = s$, alors nous obtenons encore la bonne majoration. Par contre, si nous fixons $A_s = \log(s+1)$, alors

$$\begin{aligned} \int_t^{+\infty} e^{-\alpha^2(s-t)-\alpha(A_s-A_t)} ds &= e^{\alpha^2 t}(t+1)^\alpha \int_t^{+\infty} \frac{e^{-\alpha^2 s}}{(s+1)^\alpha} ds \\ &= \frac{e^{\alpha^2 t}(t+1)^\alpha \alpha^{2\alpha}}{\alpha^2} \int_{\alpha^2 t}^{+\infty} \frac{e^{-u}}{(u+\alpha^2)^\alpha} du. \end{aligned}$$

Or le théorème de convergence dominée nous assure que

$$\lim_{\alpha \rightarrow 0} \int_{\alpha^2 t}^{+\infty} \frac{e^{-u}}{(u+\alpha^2)^\alpha} du = \int_0^{+\infty} e^{-u} du = 1.$$

Finalement, nous avons, lorsque $\alpha \rightarrow 0$,

$$\int_t^{+\infty} e^{-\alpha^2(s-t)-\alpha(A_s-A_t)} ds \sim \frac{1}{\alpha^2}.$$

Donc, ce cas particulier nous montre que l'inégalité $|\alpha Y^\alpha| \leq C$ n'est pas nécessairement vérifiée. Pour obtenir une telle inégalité dans notre cas, il apparaît nécessaire d'étudier plus finement le comportement de $(K_t^x)_{t \geq 0}$. On pourrait notamment penser à utiliser les estimations que nous avons obtenues au chapitre précédent montrant qu'en moyenne $(K_t^x)_{t \geq 0}$ se comporte comme $t \mapsto Ct$ (cf. les démonstrations des théorèmes 2.15 et 2.20).

Le second résultat nécessaire peut s'obtenir, dans le cas classique, en comparant deux solutions $Y^{x,\alpha}$ et $Y^{y,\alpha}$. Si nous regardons l'EDSR vérifiée par $Y^{x,\alpha} - Y^{y,\alpha}$, il apparaît dans notre cas le terme supplémentaire

$$\left| \int_t^T [g(X_s^x) - \alpha Y_s^{x,\alpha}] dK_s^x - \int_t^T [g(X_s^y) - \alpha Y_s^{y,\alpha}] dK_s^y \right|,$$

qu'il faudrait pouvoir majorer uniformément en α par $C|x-y|$. Nous n'avons malheureusement pas réussi à traiter ce problème directement. L'idée qui consiste à approcher l'EDS réfléchie par une EDS pénalisée et ensuite utiliser un résultat de convergence pour les EDSRs - voir par exemple l'article [18] - ne semble pas non plus aboutir.

3.2 Un nouveau théorème d'existence pour les EDSREs lorsque le générateur est non borné

Nous allons voir dans ce paragraphe comment il est possible d'obtenir un théorème d'existence et d'unicité pour l'EDSRE lorsque λ est fixé et μ fait partie des inconnues du même type que le théorème 2.20. Rappelons que ce théorème est obtenu dans le cas où le générateur ψ est non borné ce qui est le cas intéressant lorsque l'on s'intéresse aux applications aux problèmes de contrôle ergodique. La clef de cette démonstration repose sur une inégalité de grande déviation pour la moyenne empirique $\frac{1}{T} \int_0^T \mathcal{L}\phi(X_s) ds$ démontrée dans l'article [52]. Pour pouvoir appliquer cette inégalité, nous avons besoin de supposer que le processus progressif X vérifie une inégalité de Poincaré de constante c_P . Afin d'obtenir des hypothèses vérifiables en pratique, nous nous sommes cantonnés au cas des processus de Kolmogorov pour lesquels le critère de Bakry-Émery nous donne explicitement la constante c_P . Néanmoins il serait tout à fait possible de mener les mêmes calculs dans le cas général avec une constante c_P non explicite.

Dans l'article [38], les auteurs démontrent la même inégalité de grande déviation pour la moyenne empirique $\frac{1}{T} \int_0^T \mathcal{L}\phi(X_s) ds$ sans supposer que le processus progressif vérifie une inégalité de Poincaré mais en utilisant le fait que ce processus soit strictement dissipatif ce qui rentre parfaitement dans notre cadre d'étude. Il est alors possible d'obtenir un théorème d'existence et d'unicité équivalent au théorème 2.20 en remplaçant l'hypothèse (F2'') par une hypothèse ad hoc. Par soucis de lisibilité, nous allons refaire la démonstration dans son intégralité après avoir introduit notre nouvelle hypothèse (F2''') et énoncé le résultat. Toutes les notations sont celles du chapitre précédent.

$$(F2'''). \quad \left(\frac{K_{\mathcal{L}\phi} |\eta|}{|\sigma|_{\infty, \overline{G}}} + |{}^t \nabla \phi \sigma|_{\infty, \overline{G}} \right) K_{\psi, z} < -\mathbb{E}_\nu[\mathcal{L}\phi],$$

avec η la constante de dissipativité du processus X et $K_{\mathcal{L}\phi}$ la constante de lipschitz de $\mathcal{L}\phi$.

Théorème 3.2 (existence et unicité d'une solution) *supposons que les hypothèses (G1), (G2), (G3), (G4), (H1), (H2), (H3), (F1) et (F2''') sont vérifiées. Alors pour tout $\lambda \in \mathbb{R}$ il existe $\mu \in \mathbb{R}$, $v \in \mathcal{C}_{lip}^0(\overline{G})$ et une fonction mesurable $\zeta : \mathbb{R}^d \rightarrow \mathbb{R}^{1 \times d}$ tels que, si l'on définit $Y_t^x := v(X_t^x)$ et $Z_t^x := \zeta(X_t^x)$ alors $Z^x \in \mathcal{M}^2(\mathbb{R}^+, \mathbb{R}^{1 \times d})$ et \mathbb{P} -p.s. (Y^x, Z^x, μ) est une solution de l'EDSRE (2.13) avec λ fixé, pour tout $x \in \overline{G}$. De plus μ est unique parmi les solutions (Y, Z, μ) , avec λ fixé, telles que Y est un processus adapté borné continu et $Z \in \mathcal{M}^2(\mathbb{R}^+, \mathbb{R}^{1 \times d})$.*

Preuve. Comme nous l'avons déjà vu dans la preuve du théorème 2.15, il est suffisant de prouver qu'il existe une constante $C > 0$ et $x \in \overline{G}$ telles que $\mathbb{E}^{\mathbb{Q}_T} \left[\frac{K_{\mathcal{L}\phi}^x}{T} \right] \geq C$ pour tout $\mu \neq \tilde{\mu}$. Grâce à (F2''') il existe ε dans l'intervalle

$$\left[\frac{K_{\mathcal{L}\phi} |\eta|}{|\sigma|_{\infty, \overline{G}}} K_{\psi, z}, -\mathbb{E}_\nu[\mathcal{L}\phi] - |{}^t \nabla \phi \sigma|_{\infty, \overline{G}} K_{\psi, z} \right].$$

Considérons l'ensemble

$$A_T := \left\{ -\frac{1}{T} \int_0^T \mathcal{L}\phi(X_s^x) ds < -\mathbb{E}_\nu[\mathcal{L}\phi] - \varepsilon \right\} \subset \Omega,$$

avec $T > 0$ et $x \in \overline{G}$ quelconque. Nous avons

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}_T} \left[\frac{K_{\mathcal{L}\phi}^x}{T} \right] &= \mathbb{E}^{\mathbb{Q}_T} \left[\frac{\phi(X_T^x)}{T} - \frac{\phi(x)}{T} - \frac{1}{T} \int_0^T \mathcal{L}\phi(X_s^x) ds - \frac{1}{T} \int_0^T {}^t \nabla \phi(X_s^x) \sigma(X_s^x) \beta_s ds \right] \\ &\geq -\frac{2|\phi|_{\infty, \overline{G}}}{T} + \mathbb{E}^{\mathbb{Q}_T} \left[(-\mathbb{E}_\nu[\mathcal{L}\phi] - \varepsilon) \mathbb{1}_{A_T} - |\mathcal{L}\phi|_{\infty, \overline{G}} \mathbb{1}_{A_T} \right] - |{}^t \nabla \phi \sigma|_{\infty, \overline{G}} K_{\psi, z} \\ &\geq -\frac{2|\phi|_{\infty, \overline{G}}}{T} + (-\mathbb{E}_\nu[\mathcal{L}\phi] - \varepsilon) (1 - \mathbb{Q}_T(A_T)) - |\mathcal{L}\phi|_{\infty, \overline{G}} \mathbb{Q}_T(A_T) - |{}^t \nabla \phi \sigma|_{\infty, \overline{G}} K_{\psi, z}. \end{aligned}$$

En utilisant l'inégalité de Hölder avec $p > 1$ et $q > 1$ tels que $1/p + 1/q = 1$, nous obtenons

$$\begin{aligned} \mathbb{Q}_T(A_T) &= \mathbb{E} \left[\exp \left(\int_0^T \beta_s dW_s - \frac{1}{2} \int_0^T |\beta_s|^2 ds \right) \mathbb{1}_{A_T} \right] \\ &\leq \mathbb{E} \left[\exp \left(p \int_0^T \beta_s dW_s - \frac{p^2}{2} \int_0^T |\beta_s|^2 ds + \frac{p(p-1)}{2} \int_0^T |\beta_s|^2 ds \right) \right]^{1/p} \mathbb{P}_x(A_T)^{1/q} \\ &\leq \exp \left(\frac{(p-1)}{2} K_{\psi, z}^2 T \right) \mathbb{P}(A_T)^{1-1/p}. \end{aligned}$$

Pour conclure nous allons utiliser la proposition suivante que nous prouverons plus tard en utilisant le corollaire 5.11 de l'article [38] :

Proposition 3.1 *Supposons que les hypothèses (G1), (G2), (G3), (G4) et (H1) sont vérifiées et que $\eta < 0$. Alors il existe $C > 0$ telle que*

$$\mathbb{P}(A_T) \leq \exp \left(-\frac{|\sigma|_{\infty, \overline{G}}^2 \left(\varepsilon - \frac{C}{T} \right)^2 T}{2K_{\mathcal{L}\phi}^2 \eta^2} \right).$$

Ainsi,

$$\mathbb{Q}_T(A_T) \leq \exp \left[\underbrace{\left(pK_{\psi, z}^2 - \frac{|\sigma|_{\infty, \overline{G}}^2 \left(\varepsilon - \frac{C}{T} \right)^2}{K_{\mathcal{L}\phi}^2 \eta^2} \right)}_{B_p} \frac{(p-1)T}{2p} \right].$$

Nous avons choisi $\varepsilon > K_{\mathcal{L}\phi} |\eta| K_{\psi,z} / |\sigma|_{\infty, \overline{G}}$, donc il existe $T_0 \geq 0$ et une constante $\alpha > 0$ tels que pour tout $T \geq T_0$, on a $B_1 < -\alpha$. Nous pouvons donc trouver $p > 1$ et une constante $\tilde{\alpha} > 0$ tels que $B_p < -\tilde{\alpha}$ pour tout $T \geq T_0$. Finalement, nous avons prouvé que $\mathbb{Q}_T(A_T) \xrightarrow{T \rightarrow +\infty} 0$ et

$$\lim_{T \rightarrow +\infty} \mathbb{E}^{\mathbb{Q}_T} \left[\frac{K_T^x}{T} \right] \geq -\mathbb{E}_\nu[\mathcal{L}\phi] - |\nabla \phi \sigma|_{\infty, \overline{G}} K_{\psi,z} - \varepsilon > 0.$$

□

Preuve de la proposition 3.1. La démonstration de ce résultat consiste à utiliser l'inégalité de Hoeffding du corollaire 5.11 de l'article [38]. Pour cela, nous avons besoin de réécrire l'événement A_T :

$$\begin{aligned} \mathbb{P}(A_T) &= \mathbb{P} \left(-\frac{1}{T} \int_0^T \mathcal{L}\phi(X_s^x) ds < -\mathbb{E}_\nu[\mathcal{L}\phi] - \varepsilon \right) \\ &= \mathbb{P} \left(-\frac{1}{T} \int_0^T \mathcal{L}\phi(X_s^x) ds + \frac{1}{T} \int_0^T \mathbb{E}[\mathcal{L}\phi(X_s^x)] ds < -\varepsilon - \mathbb{E}_\nu[\mathcal{L}\phi] + \frac{1}{T} \int_0^T \mathbb{E}[\mathcal{L}\phi(X_s^x)] ds \right). \end{aligned}$$

Or,

$$\begin{aligned} \left| \frac{1}{T} \int_0^T \mathbb{E}[\mathcal{L}\phi(X_s^x)] ds - \mathbb{E}_\nu[\mathcal{L}\phi] \right| &= \left| \frac{1}{T} \int_0^T \mathbb{E}[\mathcal{L}\phi(X_s^x)] ds - \frac{1}{T} \int_0^T \int_{\overline{G}} \mathbb{E}[\mathcal{L}\phi(X_s^y)] \nu(dy) ds \right| \\ &\leq \frac{C}{T} \int_0^T \int_{\overline{G}} \mathbb{E}[|X_s^x - X_s^y|] \nu(dy) ds \\ &\leq \frac{C}{T} \int_0^T \int_{\overline{G}} e^{\eta s} |x - y| \nu(dy) ds \\ &\leq \frac{C}{T}, \end{aligned}$$

ce qui nous donne

$$\mathbb{P}(A_T) \leq \mathbb{P} \left(-\frac{1}{T} \int_0^T \mathcal{L}\phi(X_s^x) ds + \frac{1}{T} \int_0^T \mathbb{E}[\mathcal{L}\phi(X_s^x)] ds < -\varepsilon + \frac{C}{T} \right).$$

Le résultat de [38] permet d'estimer le membre de droite lorsque X est solution d'une EDS non réfléchie. Pour obtenir le résultat nous allons donc introduire le processus pénalisé $X^{n,x}$ solution de l'EDS

$$X_t^{n,x} = x + \int_0^t b(X_s^{n,x}) ds - \int_0^t n d(X_s^{n,x}, \overline{G}) ds + \int_0^t \sigma(X_s^{n,x}) dW_s,$$

avec $d(\cdot, \overline{G})$ la distance au compact \overline{G} . Pour n suffisamment grand, $X^{n,x}$ a pour constante de dissipativité η . De plus, nous pouvons définir σ sur tout \mathbb{R}^d en considérant $\sigma(d(\cdot, \overline{G}))$ ce qui implique que $|\sigma|_\infty = |\sigma|_{\infty, \overline{G}}$. Ainsi, l'inégalité de Hoeffding du corollaire 5.11 de l'article [38] nous assure que, pour n suffisamment grand,

$$\mathbb{P} \left(-\frac{1}{T} \int_0^T \mathcal{L}\phi(X_s^{n,x}) ds + \frac{1}{T} \int_0^T \mathbb{E}[\mathcal{L}\phi(X_s^{n,x})] ds \leq -\varepsilon + \frac{C}{T} \right) \leq \exp \left(-\frac{|\sigma|_{\infty, \overline{G}}^2 \left(\varepsilon - \frac{C}{T} \right)^2 T}{2K_{\mathcal{L}\phi}^2 \eta^2} \right).$$

Or, l'article [69] nous donne le résultat de convergence

$$\mathbb{E} \left[\sup_{s \in [0, T]} |X_s^{n,x} - X_s^x| \right] \xrightarrow{n \rightarrow +\infty} 0,$$

Ce qui nous permet de montrer, en utilisant le caractère lipschitzien de $\mathcal{L}\phi$,

$$-\frac{1}{T} \int_0^T \mathcal{L}\phi(X_s^{n,x}) ds + \frac{1}{T} \int_0^T \mathbb{E}[\mathcal{L}\phi(X_s^{n,x})] ds \xrightarrow{L^1} -\frac{1}{T} \int_0^T \mathcal{L}\phi(X_s^x) ds + \frac{1}{T} \int_0^T \mathbb{E}[\mathcal{L}\phi(X_s^x)] ds$$

lorsque $n \rightarrow +\infty$. Comme la convergence dans L^1 implique la convergence en loi, nous obtenons

$$\begin{aligned} \mathbb{P}(A_T) &\leq \liminf_{n \rightarrow +\infty} \mathbb{P} \left(-\frac{1}{T} \int_0^T \mathcal{L}\phi(X_s^{n,x}) ds + \frac{1}{T} \int_0^T \mathbb{E} [\mathcal{L}\phi(X_s^{n,x})] ds \leq -\varepsilon + \frac{C}{T} \right) \\ &\leq \exp \left(-\frac{|\sigma|_{\infty, \overline{G}}^2 \left(\varepsilon - \frac{C}{T} \right)^2 T}{2K_{\mathcal{L}\phi}^2 \eta^2} \right). \end{aligned}$$

□

Comme dans le chapitre précédent, ce résultat d'existence permet d'obtenir des résultats concernant les EDPs ergodiques et les problèmes de contrôle ergodique. Concernant les EDPs ergodiques, le théorème qui suit se démontre comme le théorème 2.24.

Théorème 3.3 *Supposons que les hypothèses du théorème 3.2 sont vérifiées. Pour tout $\lambda \in \mathbb{R}$, il existe une solution $(v(X.), \zeta(X.), \mu)$ de l'EDSRE (2.9) avec λ fixé. Alors $(v, \mu) \in \mathcal{C}_{lip}^0(\overline{G}) \times \mathbb{R}$ est une solution de viscosité de l'EDP elliptique (2.20) avec λ fixé.*

Pour le problème de contrôle ergodique, le théorème suivant se démontre comme le théorème 2.29.

Théorème 3.4 *Supposons que les hypothèses du théorème 2.15 sont vérifiées. Soit (Y, Z, μ) la solution de l'EDSRE (2.13) avec λ fixé. Alors nous avons les résultats suivants :*

1. *Pour tout contrôle ρ nous avons $J(x, \rho) \geq \mu$ et l'égalité est vérifiée si $L(X_t^x, \rho_t) + Z_t^x R(\rho_t) = \psi(X_t^x, Z_t^x)$, \mathbb{P} -p.s. pour presque tout t .*
2. *Si l'infimum est atteint dans (2.24) alors le contrôle $\overline{\rho}_t = \gamma(X_t^x, Z_t)$ vérifie $J(x, \overline{\rho}) = \mu$.*

Deuxième partie

Unicité des EDSRs quadratiques

Chapitre 4

Unicité des solutions d'EDSRs quadratiques dont le générateur est convexe et la condition terminale non bornée

Résumé : Les auteurs de l'article [25] ont prouvé un résultat d'unicité pour les solutions d'EDSRs quadratiques de générateur convexe et de condition terminale non bornée ayant tous leurs moments exponentiels finis. Dans ce papier, nous prouvons que ce résultat d'unicité reste vrai pour des solutions qui admettent uniquement certains moments exponentiels finis. Ces moments exponentiels sont reliés de manière naturelle à ceux présents dans le théorème d'existence. À l'aide de ce résultat d'unicité nous pouvons améliorer la formule de Feynman-Kac non linéaire prouvée dans [25].

Mots clés. Équations différentielles stochastiques rétrogrades, générateur à croissance quadratique, condition terminale non bornée, résultat d'unicité, formule de Feynman-Kac non linéaire.

Abstract: In [25], the authors proved the uniqueness among the solutions of quadratic BSDEs with convex generators and unbounded terminal conditions which admit every exponential moments. In this paper, we prove that uniqueness holds among solutions which admit some given exponential moments. These exponential moments are natural as they are given by the existence theorem. Thanks to this uniqueness result we can strengthen the nonlinear Feynman-Kac formula proved in [25].

Key words. Backward stochastic differential equations, generator of quadratic growth, unbounded terminal condition, uniqueness result, nonlinear Feynman-Kac formula.

AMS subject classifications. 60H10.

This chapter will be published in
Annales de l'institut Henri Poincaré
under the title:

On the uniqueness of solutions to quadratic BSDEs with convex generators and unbounded terminal conditions.

It has been written in collaboration with Freddy Delbaen and Ying Hu.

4.1 Introduction

In this paper, we consider the following quadratic backward stochastic differential equation (BSDE in short for the remaining of the paper)

$$Y_t = \xi - \int_t^T g(s, Y_s, Z_s) ds + \int_t^T Z_s dW_s, \quad 0 \leq t \leq T, \quad (4.1)$$

where the generator $-g$ is a continuous real function that is concave and has a quadratic growth with respect to the variable z . Moreover ξ is an unbounded random variable (see e.g. [60] for the case of quadratic BSDEs with bounded terminal conditions). Let us recall that, in the previous equation, we are looking for a pair of processes (Y, Z) which is required to be adapted with respect to the filtration generated by the \mathbb{R}^d -valued Brownian motion W . In [25], the authors prove the uniqueness among the solutions which satisfy for any $p > 0$,

$$\mathbb{E} \left[e^{p \sup_{0 \leq t \leq T} |Y_t|} \right] < \infty.$$

The main contribution of this paper is to strengthen their uniqueness result. More precisely, we prove the uniqueness among the solutions satisfying:

$$\exists p > \bar{\gamma}, \exists \varepsilon > 0, \quad \mathbb{E} \left[e^{p \sup_{0 \leq t \leq T} (Y_t^- + \int_0^t \bar{\alpha}_s ds)} + e^{\varepsilon \sup_{0 \leq t \leq T} Y_t^+} \right] < +\infty,$$

where $\bar{\gamma} > 0$ and $(\bar{\alpha}_t)_{t \in [0, T]}$ is a progressively measurable nonnegative stochastic process such that, \mathbb{P} -a.s.,

$$\forall (t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^{1 \times d}, \quad g(t, y, z) \leq \bar{\alpha}_t + \bar{\beta} |y| + \frac{\bar{\gamma}}{2} |z|^2.$$

Our method is different from that in [25] where the authors apply the so-called θ -difference method, i.e. estimating $Y^1 - \theta Y^2$, for $\theta \in (0, 1)$, and then letting $\theta \rightarrow 0$. Whereas in this paper, we apply a verification method: first we define a stochastic control problem and then we prove that the first component of any solution of the BSDE is the optimal value of this associated control problem. Thus the uniqueness follows immediately. Moreover, using this representation, we are able to give a probabilistic representation of the following PDE:

$$\partial_t u(t, x) + \mathcal{L}u(t, x) - g(t, x, u(t, x), -\sigma^* \nabla_x u(t, x)) = 0, \quad u(T, \cdot) = h,$$

where h and g have a “not too high” quadratic growth with respect to the variable x . We remark that the probabilistic representation is also given by [25] under the condition that h and g are subquadratic, i.e.:

$$\forall (t, x, y, z) \in [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^{1 \times d}, \quad |h(x)| + |g(t, x, y, z)| \leq f(t, y, z) + C |x|^p$$

with $f \geq 0$, $C > 0$ and $p < 2$.

The paper is organized as follows. In section 2, we prove an existence result in the spirit of [24] and [25]: here we work with generators $-g$ such that g^- has a linear growth with respect to variables y and z . As in part 5 of [24], this assumption allows us to reduce hypothesis of [25]. Section 3 is devoted to the optimal control problem from which we get as a byproduct a uniqueness result for quadratic BSDEs with unbounded terminal conditions. Finally, in the last section we derive the nonlinear Feynman-Kac formula in this framework.

Let us close this introduction by giving the notations that we will use in all the paper. For the remaining of the paper, let us fix a nonnegative real number $T > 0$. First of all, $(W_t)_{t \in [0, T]}$ is a standard Brownian motion with values in \mathbb{R}^d defined on some complete probability space $(\Omega, \mathcal{F}, \mathbb{P})$. $(\mathcal{F}_t)_{t \geq 0}$ is the natural filtration of the Brownian motion W augmented by the \mathbb{P} -null sets of \mathcal{F} . The sigma-field of predictable subsets of $[0, T] \times \Omega$ is denoted by \mathcal{P} .

As mentioned before, we will deal only with real valued BSDEs which are equations of type (4.1). The function $-g$ is called the generator and ξ the terminal condition. Let us recall that a generator is a random function $[0, T] \times \Omega \times \mathbb{R} \times \mathbb{R}^{1 \times d} \rightarrow \mathbb{R}$ which is measurable with respect to $\mathcal{P} \otimes \mathcal{B}(\mathbb{R}) \otimes \mathcal{B}(\mathbb{R}^{1 \times d})$ and a terminal condition is simply a real \mathcal{F}_T -measurable random variable. By a solution to the BSDE (4.1) we mean a pair $(Y_t, Z_t)_{t \in [0, T]}$ of predictable processes with values in $\mathbb{R} \times \mathbb{R}^{1 \times d}$ such that \mathbb{P} -a.s., $t \mapsto Y_t$ is continuous, $t \mapsto Z_t$ belongs to $L^2(0, T)$, $t \mapsto g(t, Y_t, Z_t)$ belongs to $L^1(0, T)$ and \mathbb{P} -a.s. (Y, Z) verifies

(4.1). We will sometimes use the notation $\text{BSDE}(\xi, f)$ to say that we consider the BSDE whose generator is f and whose terminal condition is ξ .

For any real $p \geq 1$, \mathcal{S}^p denotes the set of real-valued, adapted and càdlàg processes $(Y_t)_{t \in [0, T]}$ such that

$$\|Y\|_{\mathcal{S}^p} := \mathbb{E} \left[\sup_{0 \leq t \leq T} |Y_t|^p \right]^{1/p} < +\infty.$$

M^p denotes the set of (equivalent class of) predictable processes $(Z_t)_{t \in [0, T]}$ with values in $\mathbb{R}^{1 \times d}$ such that

$$\|Z\|_{M^p} := \mathbb{E} \left[\left(\int_0^T |Z_s|^2 ds \right)^{p/2} \right]^{1/p} < +\infty.$$

Finally, we will use the notation $Y^* := \sup_{0 \leq t \leq T} |Y_t|$ and we recall that Y belongs to the class (D) as soon as the family $\{Y_\tau : \tau \leq T \text{ stopping time}\}$ is uniformly integrable.

4.2 An existence result

In this section, we prove a slight modification of the existence result for quadratic BSDEs obtained in [25] by using a method applied in section 5 of [24]. We consider here the case where g^- has a linear growth with respect to variables y and z . Let us assume the following on the generator.

Assumption (A.1). There exist three constants $\beta \geq 0$, $\gamma > 0$ and $r \geq 0$ together with two progressively measurable nonnegative stochastic processes $(\bar{\alpha}_t)_{0 \leq t \leq T}$, $(\underline{\alpha}_t)_{0 \leq t \leq T}$ and a deterministic continuous nondecreasing function $\phi : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ with $\phi(0) = 0$ such that, \mathbb{P} -a.s.,

1. for all $t \in [0, T]$, $(y, z) \mapsto g(t, y, z)$ is continuous;
2. monotonicity in y : $\forall (t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^{1 \times d}$,

$$y(g(t, 0, z) - g(t, y, z)) \leq \beta |y|^2;$$

3. growth condition: $\forall (t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^{1 \times d}$,

$$-\underline{\alpha}_t - r(|y| + |z|) \leq g(t, y, z) \leq \bar{\alpha}_t + \phi(|y|) + \frac{\gamma}{2} |z|^2.$$

Theorem 4.1 *Let (A.1) hold. If there exists $p > 1$ such that*

$$\mathbb{E} \left[\exp \left(\gamma e^{\beta T} \xi^- + \gamma \int_0^T \bar{\alpha}_t e^{\beta t} dt \right) + (\xi^+)^p + \left(\int_0^T \underline{\alpha}_t dt \right)^p \right] < +\infty$$

then the BSDE (4.1) has a solution (Y, Z) such that

$$-\frac{1}{\gamma} \log \mathbb{E} \left[\exp \left(\gamma e^{\beta(T-t)} \xi^- + \gamma \int_t^T \bar{\alpha}_r e^{\beta(r-t)} dr \right) \middle| \mathcal{F}_t \right] \leq Y_t \leq C e^{CT} \left(\mathbb{E} \left[(\xi^+)^p + \left(\int_t^T \underline{\alpha}_r dr \right)^p \middle| \mathcal{F}_t \right] \right)^{1/p},$$

with C a constant that does not depend on T .

Proof. We will fit the proof of Proposition 4 in [24] to our situation. Without loss of generality, let us assume that r is an integer. For each integer $n \geq r$, let us consider the function

$$g_n(t, y, z) := \inf \{ g(t, p, q) + n|p - y| + n|q - z|, (p, q) \in \mathbb{Q}^{1+d} \}.$$

g_n is well defined and it is globally Lipschitz continuous with constant n . Moreover $(g_n)_{n \geq r}$ is increasing and converges pointwise to g . Dini's theorem implies that the convergence is also uniform on compact sets. We have also, for all $n \geq r$,

$$h(t, y, z) := -\underline{\alpha}_t - r(|y| + |z|) \leq g_n(t, y, z) \leq g(t, y, z).$$

Let (Y^n, Z^n) be the unique solution in $\mathcal{S}^p \times M^p$ to BSDE $(\xi, -g_n)$. It follows from the classical comparison theorem that

$$Y^{n+1} \leq Y^n \leq Y^r.$$

Let us prove that for each $n \geq r$

$$Y_t^n \geq -\frac{1}{\gamma} \log \mathbb{E} \left[\exp \left(\gamma e^{\beta(T-t)} \xi^- + \gamma \int_t^T \bar{\alpha}_r e^{\beta(r-t)} dr \right) \middle| \mathcal{F}_t \right] := X_t.$$

Let $(\tilde{Y}^n, \tilde{Z}^n)$ be the unique solution in $\mathcal{S}^p \times M^p$ to BSDE $(-\xi^-, -g_n^+)$. It follows from the classical comparison theorem that $\tilde{Y}^n \leq Y^n$ and $\tilde{Y}^n \leq 0$. Then, according to Proposition 3 in [25], we have $\tilde{Y}^n \geq X$ and so $Y^n \geq X$ for all $n \geq r$. We set $Y = \inf_{n \geq r} Y^n$ and, arguing as in the proof of Proposition 3 in [25] or Theorem 2 in [24] with a localization argument, we construct a process Z such that (Y, Z) is a solution to BSDE $(\xi, -g)$. For the upper bound, let (\bar{Y}, \bar{Z}) be the unique solution in $\mathcal{S}^p \times M^p$ to BSDE $(\xi^+, -h)$. Then the classical comparison theorem gives us that $Y \leq Y^n \leq \bar{Y}$ and we apply a classical a priori estimate for L^p solutions of BSDEs in [21] to \bar{Y} . \square

Corollary 4.2 *Let (A.1) hold. We suppose that $\xi^- + \int_0^T \bar{\alpha}_t dt$ has an exponential moment of order $\gamma e^{\beta T}$ and there exists $p > 1$ such that $\xi^+ + \int_0^T \underline{\alpha}_t dt \in L^p$.*

- *If $\xi^- + \int_0^T \bar{\alpha}_t dt$ has an exponential moment of order $q e^{\beta T}$ with $q > \gamma$ then the BSDE (4.1) has a solution (Y, Z) such that $\mathbb{E} [e^{qA^*}] < +\infty$ with $A_t := Y_t^- + \int_0^t \bar{\alpha}_s ds$.*
- *If $\xi^+ + \int_0^T \underline{\alpha}_t dt$ has an exponential moment of order $\varepsilon C e^{CT}$, with C given in theorem 4.1, then the BSDE (4.1) has a solution (Y, Z) such that $\mathbb{E} [e^{\varepsilon(Y^+)^*}] < +\infty$.*

Proof. Let us apply the existence result : BSDE (4.1) has a solution (Y, Z) and we have

$$A_t = Y_t^- + \int_0^t \bar{\alpha}_s ds \leq \frac{1}{\gamma} \log \mathbb{E} \left[\underbrace{\exp \left(\gamma e^{\beta T} \left(\xi^- + \int_0^T \bar{\alpha}_r dr \right) \right)}_{:= M_t} \middle| \mathcal{F}_t \right].$$

So $e^{qA_t} \leq (M_t)^{q/\gamma}$ with $q/\gamma > 1$. Since $M^{q/\gamma}$ is a submartingale, we are able to apply the Doob's maximal inequality to obtain

$$\mathbb{E} [e^{qA^*}] \leq C_q \mathbb{E} [e^{q e^{\beta T} (\xi^- + \int_0^T \bar{\alpha}_s ds)}] < +\infty.$$

To prove the second part of the corollary, we define

$$N_t := \mathbb{E} \left[(\xi^+)^p + \left| \int_0^T \underline{\alpha}_s ds \right|^p \middle| \mathcal{F}_t \right].$$

We set $q > 1$. There exists $C_{C, \varepsilon, p, q} \geq 0$ such that $x \mapsto e^{C e^{CT} x^{1/p} \varepsilon / q}$ is convex on $[C_{C, \varepsilon, p, q}, +\infty[$. We have $e^{\varepsilon/q Y_t^+} \leq e^{C e^{CT} (C_{C, \varepsilon, p, q} + N_t)^{1/p} \varepsilon / q}$. Since $e^{(C_{C, \varepsilon, p, q} + N)^{1/p} \varepsilon / q}$ is a submartingale, we are able to apply the Doob's maximal inequality to obtain

$$\mathbb{E} [e^{\varepsilon(Y^+)^*}] \leq C \mathbb{E} [e^{C e^{CT} (C_{C, \varepsilon, p, q} + (\xi^+)^p + (\int_0^T \underline{\alpha}_s ds)^p)^{1/p} \varepsilon}] \leq C \mathbb{E} [e^{\varepsilon C e^{CT} (\xi^+ + \int_0^T \underline{\alpha}_s ds)}] < +\infty.$$

\square

4.3 A uniqueness result

To prove our uniqueness result for the BSDE (4.1), we will introduce a stochastic control problem. For this purpose, we use the following assumption on g :

Assumption (A.2). There exist three constants $K_{g,y} \geq 0$, $\bar{\beta} \geq 0$ and $\bar{\gamma} > 0$ together with a progressively measurable nonnegative stochastic process $(\bar{\alpha}_t)_{t \in [0,T]}$ such that, \mathbb{P} -a.s.,

- for each $(t, z) \in [0, T] \times \mathbb{R}^{1 \times d}$,

$$|g(t, y, z) - g(t, y', z)| \leq K_{g,y} |y - y'|, \quad \forall (y, y') \in \mathbb{R}^2,$$

- for each $(t, y, z) \in [0, T] \times \mathbb{R} \times \mathbb{R}^{1 \times d}$,

$$g(t, y, z) \leq \bar{\alpha}_t + \bar{\beta} |y| + \frac{\bar{\gamma}}{2} |z|^2,$$

- for each $(t, y) \in [0, T] \times \mathbb{R}$, $z \mapsto g(t, y, z)$ is a convex function.

Since $g(t, y, \cdot)$ is a convex function we can define the Legendre-Fenchel transformation of g :

$$f(t, y, q) := \sup_{z \in \mathbb{R}^{1 \times d}} (zq - g(t, y, z)), \quad \forall t \in [0, T], q \in \mathbb{R}^d, y \in \mathbb{R}.$$

f is a function with values in $\mathbb{R} \cup \{+\infty\}$ that verifies direct properties.

Proposition 4.3

- $\forall (t, y, y', q) \in [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}^d$ such that $f(t, y, q) < +\infty$,

$$f(t, y', q) < +\infty \quad \text{and} \quad |f(t, y, z) - f(t, y', z)| \leq K_{g,y} |y - y'|.$$

- f is a convex function in q ,
- $f(t, y, q) \geq -\bar{\alpha}_t - \bar{\beta} |y| + \frac{1}{2\bar{\gamma}} |q|^2$.

For $\varepsilon > 0$ and $p > \bar{\gamma}$ given we set $N \in \mathbb{N}^*$ such that

$$\frac{T}{N} < \left(\frac{1}{\bar{\gamma}} - \frac{1}{p} \right) \frac{1}{\bar{\beta}(1/p + 1/\varepsilon)}. \quad (4.2)$$

For $i \in \{0, \dots, N\}$ we define $t_i := \frac{iT}{N}$ and, for all real $\mathcal{F}_{t_{i+1}}$ -measurable random variable η ,

$$\begin{aligned} \mathcal{A}_{t_i, t_{i+1}}(\eta) &:= \left\{ (q_s)_{s \in [t_i, t_{i+1}]} : \int_{t_i}^{t_{i+1}} |q_s|^2 ds < +\infty \text{ } \mathbb{P}\text{-a.s.}, \quad \mathbb{E}^{\mathbb{Q}^i} \left[\int_{t_i}^{t_{i+1}} |q_s|^2 ds \right] < +\infty, \right. \\ &\quad (M_t^i)_{t \in [t_i, t_{i+1}]} \text{ is a martingale, } \quad \mathbb{E}^{\mathbb{Q}^i} \left[|\eta| + \int_{t_i}^{t_{i+1}} |f(s, 0, q_s)| ds \right] < +\infty, \\ &\quad \left. \text{with } M_t^i := \exp \left(\int_{t_i}^t q_s dW_s - \frac{1}{2} \int_{t_i}^t |q_s|^2 ds \right) \text{ and } \frac{d\mathbb{Q}^i}{d\mathbb{P}} := M_{t_{i+1}}^i \right\}. \end{aligned}$$

Let q be in $\mathcal{A}_{t_i, t_{i+1}}(\eta)$, if this set is not empty. We define $dW_t^q := dW_t - q_t dt$. Thanks to the Girsanov theorem, $(W_{t_i+h}^q - W_{t_i}^q)_{h \in [0, T/N]}$ is a Brownian motion under the probability \mathbb{Q}^i . So, we are able to apply Proposition 6.4 in [21] to obtain this existence result:

Proposition 4.4 *There exist two processes $(Y^{\eta, q}, Z^{\eta, q})$ such that $(Y_t^{\eta, q})_{t \in [t_i, t_{i+1}]}$ belongs to the class (D) under \mathbb{Q}^i , $\int_{t_i}^{t_{i+1}} |Z_s^{\eta, q}|^2 ds < +\infty$ \mathbb{P} -a.s., $\int_{t_i}^{t_{i+1}} |f(s, Y_s^{\eta, q}, q_s)| ds < +\infty$ \mathbb{P} -a.s. and*

$$Y_t^{\eta, q} = \eta + \int_t^{t_{i+1}} f(s, Y_s^{\eta, q}, q_s) ds + \int_t^{t_{i+1}} Z_s^{\eta, q} dW_s^q, \quad t_i \leq t \leq t_{i+1}.$$

We are now able to define the admissible control set:

$$\begin{aligned} \mathcal{A} &:= \left\{ (q_s)_{s \in [0, T]} : q_{[t_{N-1}, T]} \in \mathcal{A}_{t_{N-1}, T}(\xi), \forall i \in \{N-2, \dots, 0\}, q_{[t_i, t_{i+1}]} \in \mathcal{A}_{t_i, t_{i+1}}(Y_{t_{i+1}}^q) \right. \\ &\quad \left. \text{with } Y_{t_{i+1}}^q := Y_{t_{i+1}}^{Y_{t_{i+2}}^q, q_{[t_{i+1}, t_{i+2}]}} \text{ and } Y_T^q := \xi \right\}. \end{aligned}$$

\mathcal{A} is well defined by a decreasing recursion on $i \in \{0, \dots, N-1\}$. For $q \in \mathcal{A}$ we can define our cost functional Y^q on $[0, T]$ by

$$\forall i \in \{N-1, \dots, 0\}, \forall t \in [t_i, t_{i+1}], \quad Y_t^q := Y_{t_{i+1}}^{Y_{t_{i+1}}^q, q|_{[t_i, t_{i+1}]}} ,$$

and, similarly, we define the process Z^q associated to Y^q by

$$\forall i \in \{N-1, \dots, 0\}, \forall t \in (t_i, t_{i+1}), \quad Z_t^q := Z_{t_{i+1}}^{Y_{t_{i+1}}^q, q|_{[t_i, t_{i+1}]}} .$$

(Y^q, Z^q) is also well defined by a decreasing recursion on $i \in \{0, \dots, N-1\}$. Finally, the stochastic control problem consists in minimizing Y^q among all the admissible controls $q \in \mathcal{A}$. Our strategy to prove the uniqueness is to prove that given a solution (Y, Z) , the first component is the optimal value.

Theorem 4.5 *We suppose that there exists a solution (Y, Z) of the BSDE (4.1) verifying*

$$\exists p > \bar{\gamma}, \exists \varepsilon > 0, \quad \mathbb{E} [\exp(pA^*) + \exp(\varepsilon(Y^+)^*)] < +\infty,$$

with $A_t := Y_t^- + \int_0^t \bar{\alpha}_s ds$. Then we have $Y = \text{ess inf}_{q \in \mathcal{A}} Y^q$, and there exists $q^* \in \mathcal{A}$ such that $Y = Y^{q^*}$. Moreover, this implies that the solution (Y, Z) is unique among solutions verifying such condition.

Proof. Let us first prove that for any q admissible, we have $Y \leq Y^q$. To do this, we will show that $Y_{[t_i, t_{i+1}]} \leq Y_{[t_i, t_{i+1}]}^q$ by decreasing recurrence on $i \in \{0, \dots, N-1\}$. Firstly, we have $Y_T = Y_T^q = \xi$. Then we suppose that $Y_t \leq Y_t^q, \forall t \in [t_{i+1}, T]$. We set $t \in [t_i, t_{i+1}]$ and we define

$$\tau_n^i := \inf \left\{ s \geq t : \sup \left\{ \int_t^s |Z_u|^2 du, \int_t^s |Z_u^q|^2 du, \int_t^s |q_u|^2 du \right\} > n \right\} \wedge t_{i+1},$$

$$h(s, y, z) := -g(s, y, z) + zq_s, \text{ and}$$

$$h_s := \begin{cases} \frac{h(s, Y_s^q, Z_s) - h(s, Y_s, Z_s)}{Y_s^q - Y_s} & \text{if } Y_s^q - Y_s \neq 0 \\ 0 & \text{otherwise.} \end{cases}$$

We observe that $|h_s| \leq K_{g,y}$. Then, by applying Itô formula to the process $(Y_s^q - Y_s)e^{\int_t^s h_u du}$ we obtain

$$Y_t^q - Y_t = e^{\int_t^{\tau_n^i} h_s ds} [Y_{\tau_n^i}^q - Y_{\tau_n^i}] + \int_t^{\tau_n^i} e^{\int_t^s h_u du} [f(s, Y_s^q, q_s) - h(s, Y_s^q, Z_s)] ds + \int_t^{\tau_n^i} e^{\int_t^s h_u du} [Z_s^q - Z_s] dW_s^q.$$

By definition, $f(s, Y_s^q, q_s) - h(s, Y_s^q, Z_s) \geq 0$, so

$$Y_t^q - Y_t \geq \mathbb{E}^{\mathbb{Q}^i} \left[e^{\int_t^{\tau_n^i} h_s ds} [Y_{\tau_n^i}^q - Y_{\tau_n^i}] \middle| \mathcal{F}_t \right].$$

Since $\left(Y_{\tau_n^i}^q e^{\int_t^{\tau_n^i} h_s ds} \right)_n$ tends to $Y_{t_{i+1}}^q e^{\int_t^{t_{i+1}} h_s ds}$ almost surely and is uniformly integrable under \mathbb{Q}^i , we have

$$\lim_{n \rightarrow +\infty} \mathbb{E}^{\mathbb{Q}^i} \left[e^{\int_t^{\tau_n^i} h_s ds} Y_{\tau_n^i}^q \middle| \mathcal{F}_t \right] = \mathbb{E}^{\mathbb{Q}^i} \left[e^{\int_t^{t_{i+1}} h_s ds} Y_{t_{i+1}}^q \middle| \mathcal{F}_t \right].$$

Moreover, $\left| Y_{\tau_n^i}^q e^{\int_t^{\tau_n^i} h_s ds} \right| \leq (Y^+)^* e^{TK_{g,y}} + (Y^-)^* e^{TK_{g,y}}$. Let us recall a useful inequality: from

$$xy \leq \exp(x) + y(\log(y) - 1), \quad \forall (x, y) \in \mathbb{R} \times \mathbb{R}^+,$$

we deduce

$$xy = px \frac{y}{p} \leq \exp(px) + \frac{y}{p} (\log y - \log p - 1). \quad (4.3)$$

Thus

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}^i} [(Y^-)^*] &= \mathbb{E} [M_{t_{i+1}}^i (Y^-)^*] \leq \mathbb{E} [\exp(p(Y^-)^*)] + \frac{1}{p} \mathbb{E} [M_{t_{i+1}}^i (\log M_{t_{i+1}}^i - \log p - 1)] \\ &\leq C_p + \frac{1}{2p} \mathbb{E}^{\mathbb{Q}^i} \left[\int_{t_i}^{t_{i+1}} |q_s|^2 ds \right] \\ &< +\infty, \end{aligned}$$

and, in the same manner,

$$\mathbb{E}^{\mathbb{Q}^i} [(Y^+)^*] \leq C_\varepsilon + \frac{1}{2\varepsilon} \mathbb{E}^{\mathbb{Q}^i} \left[\int_{t_i}^{t_{i+1}} |q_s|^2 ds \right] < +\infty.$$

So, by applying the dominated convergence theorem we obtain

$$\lim_{n \rightarrow +\infty} \mathbb{E}^{\mathbb{Q}^i} \left[e^{\int_{t_i}^{\tau_n^i} h_s ds} Y_{\tau_n^i} \middle| \mathcal{F}_t \right] = \mathbb{E}^{\mathbb{Q}^i} \left[e^{\int_{t_i}^{t_{i+1}} h_s ds} Y_{t_{i+1}} \middle| \mathcal{F}_t \right].$$

Finally,

$$Y_t^q - Y_t \geq \lim_{n \rightarrow +\infty} \mathbb{E}^{\mathbb{Q}^i} \left[e^{\int_{t_i}^{\tau_n^i} h_s ds} [Y_{\tau_n^i}^q - Y_{\tau_n^i}] \middle| \mathcal{F}_t \right] = \mathbb{E}^{\mathbb{Q}^i} \left[e^{\int_{t_i}^{t_{i+1}} h_s ds} (Y_{t_{i+1}}^q - Y_{t_{i+1}}) \middle| \mathcal{F}_t \right] \geq 0,$$

because $Y_{t_{i+1}}^q \geq Y_{t_{i+1}}$ by the recurrence's hypothesis.

Now we set ${}^t q_s^* \in \partial_z g(s, Y_s, Z_s)$ with $\partial_z g(s, Y_s, Z_s)$ the subdifferential of $z \mapsto g(s, Y_s, z)$ at Z_s . We recall that for a convex function $l : \mathbb{R}^{1 \times d} \rightarrow \mathbb{R}$, the subdifferential of l at x_0 is the non-empty convex compact set of $u \in \mathbb{R}^{1 \times d}$ such that

$$l(x) - l(x_0) \geq u^t (x - x_0), \quad \forall x \in \mathbb{R}^{1 \times d}.$$

We have $f(s, Y_s, q_s^*) = Z_s q_s^* - g(s, Y_s, Z_s)$ for all $s \in [0, T]$, so

$$\begin{aligned} g(s, Y_s, Z_s) &\leq Z_s q_s^* - \frac{1}{2\bar{\gamma}} |q_s^*|^2 + \bar{\beta} |Y_s| + \bar{\alpha}_s \\ &\leq \frac{1}{2} \left(2\bar{\gamma} |Z_s|^2 + \frac{|q_s^*|^2}{2\bar{\gamma}} \right) - \frac{1}{2\bar{\gamma}} |q_s^*|^2 + \bar{\beta} |Y_s| + \bar{\alpha}_s, \\ \frac{|q_s^*|^2}{4\bar{\gamma}} &\leq -g(s, Y_s, Z_s) + \bar{\gamma} |Z_s|^2 + \bar{\beta} |Y_s| + \bar{\alpha}_s, \end{aligned}$$

and finally, $\int_0^T |q_s^*|^2 ds < +\infty$, \mathbb{P} -a.s.. Moreover, $\forall t, t' \in [0, T]$,

$$Y_t = Y_{t'} + \int_t^{t'} f(s, Y_s, q_s^*) ds + \int_t^{t'} Z_s (dW_s - q_s^* ds).$$

Thus, we just have to show that q^* is admissible to prove that q^* is optimal, i.e. $Y = Y^{q^*}$. For this, we must prove that $(q_s^*)_{s \in [t_i, t_{i+1}]} \in \mathcal{A}_{t_i, t_{i+1}}(Y_{t_{i+1}})$ for $i \in \{0, \dots, N-1\}$. We define

$$\begin{aligned} M_t^i &:= \exp \left(\int_{t_i}^t q_s^* dW_s - \frac{1}{2} \int_{t_i}^t |q_s^*|^2 ds \right), \quad \frac{dQ^{*,i}}{d\mathbb{P}} := M_{t_{i+1}}^i, \\ \tau_n^i &= \inf \left\{ t \in [t_i, t_{i+1}] : \sup \left(\int_{t_i}^t |q_s^*|^2 ds, \int_{t_i}^t |Z_s|^2 ds \right) > n \right\} \wedge t_{i+1}, \quad \frac{dQ_n^{*,i}}{d\mathbb{P}} := M_{\tau_n^i}^i. \end{aligned}$$

Let us show the following lemma:

Lemma 4.6 $(M_{\tau_n^i}^i)_n$ is uniformly integrable.

Proof. We apply inequality (4.3) to obtain

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}_n^{*,i}}[A^*] &= \mathbb{E}\left[M_{\tau_n^i}^i A^*\right] \leq \mathbb{E}[\exp(pA^*)] + \frac{1}{p}\mathbb{E}\left[M_{\tau_n^i}^i (\log M_{\tau_n^i}^i - \log p - 1)\right] \\ &\leq C_p + \frac{1}{2p}\mathbb{E}^{\mathbb{Q}_n^{*,i}}\left[\int_{t_i}^{\tau_n^i} |q_s^*|^2 ds\right],\end{aligned}$$

and, in the same manner,

$$\mathbb{E}^{\mathbb{Q}_n^{*,i}}[(Y^+)^*] \leq C_\varepsilon + \frac{1}{2\varepsilon}\mathbb{E}^{\mathbb{Q}_n^{*,i}}\left[\int_{t_i}^{\tau_n^i} |q_s^*|^2 ds\right].$$

Since $g(s, Y_s, Z_s) = Z_s q_s^* - f(s, Y_s, q_s^*)$ and $(M_{t \wedge \tau_n^i}^i)_{t \in [t_i, t_{i+1}]}$ is a martingale, we can apply the Girsanov theorem and we obtain

$$\mathbb{E}^{\mathbb{Q}_n^{*,i}}\left[Y_{\tau_n^i} + \int_{t_i}^{\tau_n^i} f(s, Y_s, q_s^*) ds\right] = \mathbb{E}^{\mathbb{Q}_n^{*,i}}[Y_{t_i}] = \mathbb{E}\left[M_{\tau_n^i}^i Y_{t_i}\right] = \mathbb{E}[Y_{t_i}].$$

Moreover $f(t, y, q) \geq \frac{1}{2\bar{\gamma}}|q|^2 - \bar{\beta}|y| - \bar{\alpha}_t$ and $Y_{\tau_n^i} \geq -Y_{\tau_n^i}^-$, so

$$\begin{aligned}\mathbb{E}[Y_{t_i}] &\geq -\mathbb{E}^{\mathbb{Q}_n^{*,i}}[Y_{\tau_n^i}^-] - \mathbb{E}^{\mathbb{Q}_n^{*,i}}\left[\int_{t_i}^{\tau_n^i} \bar{\alpha}_s ds\right] + \frac{1}{2\bar{\gamma}}\mathbb{E}^{\mathbb{Q}_n^{*,i}}\left[\int_{t_i}^{\tau_n^i} |q_s^*|^2 ds\right] - \bar{\beta}\mathbb{E}^{\mathbb{Q}_n^{*,i}}\left[\int_{t_i}^{\tau_n^i} |Y_s| ds\right] \\ &\geq C - \mathbb{E}^{\mathbb{Q}_n^{*,i}}[A^*] + \frac{1}{2\bar{\gamma}}\mathbb{E}^{\mathbb{Q}_n^{*,i}}\left[\int_{t_i}^{\tau_n^i} |q_s^*|^2 ds\right] - \frac{T}{N}\left(\bar{\beta}\mathbb{E}^{\mathbb{Q}_n^{*,i}}[(Y^-)^* + (Y^+)^*]\right) \\ &\geq C_{p,\varepsilon} + \underbrace{\frac{1}{2}\left(\frac{1}{\bar{\gamma}} - \frac{1}{p} - \frac{T}{N}\left(\frac{\bar{\beta}}{p} + \frac{\bar{\beta}}{\varepsilon}\right)\right)}_{>0}\mathbb{E}^{\mathbb{Q}_n^{*,i}}\left[\int_{t_i}^{\tau_n^i} |q_s^*|^2 ds\right].\end{aligned}$$

This inequality explains why we take N verifying (4.2). Finally we get that

$$2\mathbb{E}\left[M_{\tau_n^i}^i \log M_{\tau_n^i}^i\right] = \mathbb{E}^{\mathbb{Q}_n^{*,i}}\left[\int_{t_i}^{\tau_n^i} |q_s^*|^2 ds\right] \leq C_{p,\varepsilon}. \quad (4.4)$$

Then we conclude the proof of the lemma by using the de La Vallée Poussin lemma. \square

Thanks to this lemma, we have that $\mathbb{E}[M_{t_{i+1}}^i] = 1$ and so $(M_t^i)_{t \in [t_i, t_{i+1}]}$ is a martingale. Moreover, applying Fatou's lemma and inequality (4.4), we obtain

$$2\mathbb{E}\left[M_{t_{i+1}}^i \log M_{t_{i+1}}^i\right] = \mathbb{E}^{\mathbb{Q}_n^{*,i}}\left[\int_{t_i}^{t_{i+1}} |q_s^*|^2 ds\right] \leq \liminf_n \mathbb{E}^{\mathbb{Q}_n^{*,i}}\left[\int_{t_i}^{\tau_n^i} |q_s^*|^2 ds\right] < +\infty. \quad (4.5)$$

So, by using this result and inequality (4.3) we easily show that $\mathbb{E}^{\mathbb{Q}_n^{*,i}}[(Y^+)^* + (Y^-)^*] < +\infty$. To conclude we have to prove that $\mathbb{E}^{\mathbb{Q}_n^{*,i}}\left[\int_{t_i}^{t_{i+1}} |f(s, 0, q_s^*)| ds\right] < +\infty$:

$$\begin{aligned}\mathbb{E}^{\mathbb{Q}_n^{*,i}}\left[\int_{t_i}^{t_{i+1}} |f(s, 0, q_s^*)| ds\right] &\leq \mathbb{E}^{\mathbb{Q}_n^{*,i}}\left[\int_{t_i}^{t_{i+1}} |f(s, Y_s, q_s^*)| + K_{g,y}|Y_s| ds\right] \\ &\leq \mathbb{E}^{\mathbb{Q}_n^{*,i}}\left[\int_{t_i}^{t_{i+1}} |f(s, Y_s, q_s^*)| ds + K_{g,y}T((Y^+)^* + (Y^-)^*)\right] \\ &\leq C + \mathbb{E}^{\mathbb{Q}_n^{*,i}}\left[\int_{t_i}^{t_{i+1}} f^+(s, Y_s, q_s^*) + f^-(s, Y_s, q_s^*) ds\right].\end{aligned}$$

Firstly,

$$\mathbb{E}^{\mathbb{Q}_n^{*,i}}\left[\int_{t_i}^{t_{i+1}} f^-(s, Y_s, q_s^*) ds\right] \leq \mathbb{E}^{\mathbb{Q}_n^{*,i}}\left[\int_{t_i}^{t_{i+1}} \bar{\alpha}_s + \bar{\beta}|Y_s| ds\right] < +\infty.$$

Moreover, thanks to the Girsanov theorem we have

$$\mathbb{E}^{\mathbb{Q}^{*,i}}[Y_{t_i}] = \mathbb{E}^{\mathbb{Q}^{*,i}} \left[Y_{\tau_n^i} + \int_{t_i}^{\tau_n^i} f(s, Y_s, q_s^*) ds \right],$$

so

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}^{*,i}} \left[\int_{t_i}^{\tau_n^i} f^+(s, Y_s, q_s^*) ds \right] &\leq \mathbb{E}^{\mathbb{Q}^{*,i}}[Y_{t_i} - Y_{\tau_n^i}] + \mathbb{E}^{\mathbb{Q}^{*,i}} \left[\int_{t_i}^{\tau_n^i} f^-(s, Y_s, q_s^*) ds \right] \\ &\leq C + \mathbb{E}^{\mathbb{Q}^{*,i}} \left[\int_{t_i}^{t_{i+1}} f^-(s, Y_s, q_s^*) ds \right] \leq C \end{aligned}$$

Finally, $\mathbb{E}^{\mathbb{Q}^{*,i}} \left[\int_{t_i}^{t_{i+1}} f^+(s, Y_s, q_s^*) ds \right] < +\infty$ and $\mathbb{E}^{\mathbb{Q}^{*,i}} \left[\int_{t_i}^{t_{i+1}} |f(s, 0, q_s^*)| ds \right] < +\infty$. Thus, we prove that q^* is optimal, i.e. $Y^{q^*} = Y$.

The uniqueness of Y is a direct consequence of the fact that $Y = Y^{q^*} = \text{ess inf}_{q \in \mathcal{A}} Y^q$. The uniqueness of Z follows immediately. \square

Remark 4.7 If we have $g(t, y, z) \leq g(t, 0, z)$, then $f(t, y, q) \geq f(t, 0, q) \geq \frac{1}{2\gamma} |q|^2 - \bar{\alpha}_t$ and we do not have to introduce N in the proof of Lemma 4.6. So we have a simpler representation theorem:

$$Y_t = \text{ess inf}_{q \in \mathcal{A}_{0,T}(\xi)} Y_t^q, \quad \forall t \in [0, T].$$

For example, when g is independent of y , we obtain

$$Y_t = \text{ess inf}_{q \in \mathcal{A}_{0,T}(\xi)} \mathbb{E}^{\mathbb{Q}} \left[\xi + \int_t^T f(s, q_s) ds \middle| \mathcal{F}_t \right], \quad \forall t \in [0, T].$$

4.4 Application to quadratic PDEs

In this section we give an application of our results concerning BSDEs to PDEs which are quadratic with respect to the gradient of the solution. Let us consider the following semilinear PDE

$$\partial_t u(t, x) + \mathcal{L}u(t, x) - g(t, x, u(t, x), -\sigma^* \nabla_x u(t, x)) = 0, \quad u(T, \cdot) = h, \quad (4.6)$$

where \mathcal{L} is the infinitesimal generator of the diffusion $X^{t,x}$ solution to the SDE

$$X_s^{t,x} = x + \int_t^s b(r, X_r^{t,x}) dr + \int_t^s \sigma(r) dW_r, \quad t \leq s \leq T, \text{ and } X_s^{t,x} = x, \quad s \leq t. \quad (4.7)$$

The nonlinear Feynman-Kac formula consists in proving that the function defined by the formula

$$\forall (t, x) \in [0, T] \times \mathbb{R}^d, \quad u(t, x) := Y_t^{t,x} \quad (4.8)$$

where, for each $(t_0, x_0) \in [0, T] \times \mathbb{R}^d$, $(Y^{t_0, x_0}, Z^{t_0, x_0})$ stands for the solution to the following BSDE

$$Y_t = h(X_T^{t_0, x_0}) - \int_t^T g(s, X_s^{t_0, x_0}, Y_s, Z_s) ds + \int_t^T Z_s dW_s, \quad 0 \leq t \leq T, \quad (4.9)$$

is a solution, at least a viscosity solution, to the PDE (4.6).

Assumption (A.3). Let $b : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ and $\sigma : [0, T] \rightarrow \mathbb{R}^{d \times d}$ be continuous functions and let us assume that there exists $K \geq 0$ such that:

1. $\forall t \in [0, T], |b(t, 0)| \leq K$,
2. $\forall t \in [0, T], \forall (x, x') \in \mathbb{R}^d \times \mathbb{R}^d, |b(t, x) - b(t, x')| \leq K |x - x'|$.

Lemma 4.8

$$\forall \lambda \in \left[0, \frac{1}{2e^{2KT} \|\sigma\|_\infty^2 T} \right], \exists C_T \geq 0, \exists C \geq 0, \quad \mathbb{E} \left[\sup_{0 \leq t \leq T} e^{\lambda |X_t^{t_0, x_0}|^2} \right] \leq C_T e^{C |x_0|^2},$$

with $T \mapsto C_T$ nondecreasing.

Proof. As in [25] we easily show that, for all $\varepsilon > 0$, we have

$$\begin{aligned} \sup_{0 \leq t \leq T} |X_t^{t_0, x_0}| &\leq \left(|x_0| + KT + \sup_{0 \leq t \leq T} \left| \int_0^t \mathbb{1}_{s \geq t_0} \sigma(s) dW_s \right| \right) e^{KT} \\ \sup_{0 \leq t \leq T} |X_t^{t_0, x_0}|^2 &\leq C_\varepsilon (T^2 + |x_0|^2) + (1 + \varepsilon) e^{2KT} \sup_{0 \leq t \leq T} \left| \int_0^t \mathbb{1}_{s \geq t_0} \sigma(s) dW_s \right|^2. \end{aligned}$$

We define $\tilde{\lambda} := \lambda(1 + \varepsilon)e^{2KT}$. It follows from the Dambis-Dubins-Schwarz representation theorem and the Doob's maximal inequality that

$$\mathbb{E} \left[\sup_{0 \leq t \leq T} \exp \left(\tilde{\lambda} \left| \int_0^t \mathbb{1}_{s \geq t_0} \sigma(s) dW_s \right|^2 \right) \right] \leq \mathbb{E} \left[\sup_{0 \leq t \leq \|\sigma\|_\infty^2 T} e^{\tilde{\lambda} |W_t|^2} \right] \leq 4\mathbb{E} \left[e^{\tilde{\lambda} \|\sigma\|_\infty^2 T |W_1|^2} \right],$$

which is a finite constant if $\tilde{\lambda} \|\sigma\|_\infty^2 T < 1/2$. \square

With this observation in hands, we can give our assumptions on the nonlinear term of the PDE and the terminal condition.

Assumption (A.4). Let $g : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$ and $h : \mathbb{R}^d \rightarrow \mathbb{R}$ be continuous and let us assume moreover that there exist five constants $r \geq 0, \beta \geq 0, \gamma \geq 0, \alpha \geq 0$ and $\alpha' \geq 0$ such that:

1. for each $(t, x, z) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^{1 \times d}$,

$$\forall (y, y') \in \mathbb{R}^2, \quad |g(t, x, y, z) - g(t, x, y', z)| \leq \beta |y - y'|;$$

2. for each $(t, x, y) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}$, $z \mapsto g(t, x, y, z)$ is convex on $\mathbb{R}^{1 \times d}$;

3. for each $(t, x, y, z) \in [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^{1 \times d}$,

$$-r(1 + |x|^2 + |y| + |z|) \leq g(t, x, y, z) \leq r + \alpha |x|^2 + \beta |y| + \frac{\gamma}{2} |z|^2,$$

$$-r - \alpha' |x|^2 \leq h(x) \leq r(1 + |x|^2);$$

4. for each $(t, x, x', y, z) \in [0, T] \times \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^{1 \times d}$,

$$|g(t, x, y, z) - g(t, x', y, z)| \leq r(1 + |x| + |x'|) |x - x'|,$$

$$|h(x) - h(x')| \leq r(1 + |x| + |x'|) |x - x'|;$$

- 5.

$$\alpha' + T\alpha < \frac{1}{2\gamma e^{(2K+\beta)T} \|\sigma\|_\infty^2 T}.$$

Thanks to Lemma 4.8, we see that there exist $q > \gamma e^{\beta T}$ and $\varepsilon > 0$ such that $h^-(X_T^{t_0, x_0}) + \int_0^T \left(r + \alpha |X_t^{t_0, x_0}|^2 \right) dt$ has an exponential moment of order q and $h^+(X_T^{t_0, x_0}) + \int_0^T \left(r + r |X_t^{t_0, x_0}|^2 \right) dt$ has an exponential moment of order ε . So we are able to apply Corollary 4.2 and Theorem 4.5 to obtain a unique solution $(Y^{t_0, x_0}, Z^{t_0, x_0})$ to the BSDE (4.9). Let us recall the definition of a viscosity solution and then, prove that u is a viscosity solution to the PDE (4.6).

Definition 4.9 A continuous function u on $[0, T] \times \mathbb{R}^d$ such that $u(T, \cdot) = h$ is said to be a viscosity subsolution (respectively supersolution) to (4.6) if

$$\partial_t \varphi(t_0, x_0) + \mathcal{L}\varphi(t_0, x_0) - g(t_0, x_0, u(t_0, x_0), -\sigma^* \nabla_x \varphi(t_0, x_0)) \geq 0, \quad (\text{respectively } \leq 0)$$

as soon as $u - \varphi$ has a local maximum (respectively minimum) at $(t_0, x_0) \in (0, T) \times \mathbb{R}^d$ where φ is a smooth function. A viscosity solution is both a viscosity subsolution and a viscosity supersolution.

Proposition 4.10 *Let assumptions (A.3) and (A.4) hold. The function u defined by (4.8) is continuous on $[0, T] \times \mathbb{R}^d$ and satisfies*

$$\forall (t, x) \in [0, T] \times \mathbb{R}^d, \quad |u(t, x)| \leq C(1 + |x|^2).$$

Moreover u is a viscosity solution to the PDE (4.6).

Before giving a proof of this result, we will show some auxiliary results about admissible control sets. We have already notice in Remark 4.7 that we have a simpler representation theorem when T is small enough to take $N = 1$ in (4.2). So we define a constant $T_1 > 0$ such that for all $T \in [0, T_1]$ we are allowed to set $N = 1$. We will reuse notations of Section 4.3. For all $T \in [0, T_1]$, $t \in [0, T]$, $x \in \mathbb{R}^d$, we define the admissible control set

$$\begin{aligned} \mathcal{A}_{0,T}(t, x) := & \left\{ (q_s)_{s \in [0, T]} : \int_0^T |q_s|^2 ds < +\infty \text{ } \mathbb{P} - a.s., \quad \mathbb{E}^{\mathbb{Q}} \left[\int_0^T |q_s|^2 ds \right] < +\infty, \right. \\ & (M_t)_{t \in [0, T]} \text{ is a martingale, } \quad \mathbb{E}^{\mathbb{Q}} \left[|h(X_T^{t,x})| + \int_0^T |f(s, X_s^{t,x}, 0, q_s)| ds \right] < +\infty, \\ & \left. \text{with } M_t := \exp \left(\int_0^t q_s dW_s - \frac{1}{2} \int_0^t |q_s|^2 ds \right) \text{ and } \frac{d\mathbb{Q}}{d\mathbb{P}} := M_T \right\}. \end{aligned}$$

We will prove a first lemma and then we will use it to show that this admissible control set does not depend on t and x .

Lemma 4.11 $\exists C > 0$ such that $\forall T \in [0, T_1]$, $\forall t \in [0, T]$, $\forall x \in \mathbb{R}^d$, $\forall q \in \mathcal{A}_{0,T}(t, x)$, $\forall t' \in [0, T]$, $\forall x' \in \mathbb{R}^d$, $\forall s \in [t', T]$,

$$\mathbb{E}^{\mathbb{Q}} \left[|X_s^{t',x'}|^2 \right] \leq C \left(1 + |x'|^2 + T \int_{t'}^s \mathbb{E}^{\mathbb{Q}} [|q_u|^2] du \right).$$

Remark 4.12 *In the second part of the lemma q and \mathbb{Q} depend on x and t but we do not write it to simplify notations.*

Proof. For all $s \in [t', T]$ we have an obvious inequality

$$|X_s^{t',x'}|^2 \leq C \left(1 + |x'|^2 + \left(\int_{t'}^s |X_u^{t',x'}| du \right)^2 + \sup_{t' \leq r \leq T} \left| \int_{t'}^r \sigma(u) dW_u^q \right|^2 + \left(\int_{t'}^s |q_u| du \right)^2 \right).$$

Then, by applying Cauchy-Schwarz's inequality and Doob's maximal inequality, we obtain

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} \left[|X_s^{t',x'}|^2 \right] & \leq C \left(1 + |x'|^2 + T \int_{t'}^s \mathbb{E}^{\mathbb{Q}} \left[|X_u^{t',x'}|^2 \right] du + \mathbb{E}^{\mathbb{Q}} \left[\left| \int_{t'}^T \sigma(u) dW_u^q \right|^2 \right] \right. \\ & \quad \left. + T \mathbb{E}^{\mathbb{Q}} \left[\int_{t'}^s |q_u|^2 du \right] \right). \end{aligned}$$

Finally, Gronwall's Lemma gives us the result. \square

Proposition 4.13 $\mathcal{A}_{0,T}(t, x)$ is independent of t and x . We will write it $\mathcal{A}_{0,T}$.

Proof. Let $x, x' \in \mathbb{R}^d$, $t, t' \in [0, T]$ and $q \in \mathcal{A}_{0,T}(t, x)$. We will show that $q \in \mathcal{A}_{0,T}(t', x')$. Firstly, thanks to Lemma 4.11 we have

$$\mathbb{E}^{\mathbb{Q}} \left[|h(X_T^{t',x'})| \right] \leq C \left(1 + \mathbb{E}^{\mathbb{Q}} \left[|X_T^{t',x'}|^2 \right] \right) \leq C \left(1 + \int_{t'}^T \mathbb{E}^{\mathbb{Q}} [|q_u|^2] du \right) < +\infty.$$

Moreover

$$-C(1 + |X_s^{t',x'}|^2) \leq \frac{1}{2\gamma} |q_s|^2 - C(1 + |X_s^{t',x'}|^2) \leq f(s, X_s^{t',x'}, 0, q_s),$$

and

$$f(s, X_s^{t',x'}, 0, q_s) \leq f(s, X_s^{t,x}, 0, q_s) + C(1 + |X_s^{t,x}|^2 + |X_s^{t',x'}|^2).$$

So, $|f(s, X_s^{t',x'}, 0, q_s)| \leq |f(s, X_s^{t,x}, 0, q_s)| + C(1 + |X_s^{t,x}|^2 + |X_s^{t',x'}|^2)$ and finally

$$\mathbb{E}^{\mathbb{Q}} \left[\int_0^T |f(s, X_s^{t',x'}, 0, q_s)| ds \right] < +\infty.$$

□

Now we will do a new restriction of the admissible control set.

Proposition 4.14 $\exists T_2 \in]0, T_1], \exists \tilde{C} > 0$, such that, $\forall T \in [0, T_2], \forall t \in [0, T], \forall s \in [0, t], \forall x \in \mathbb{R}^d$,

$$|Y_s^{t,x}| \leq \tilde{C}(1 + |x|^2) \quad \text{and} \quad \mathbb{E}^{\mathbb{Q}^*} \left[\int_0^T |q_u^*|^2 du \right] \leq \tilde{C}(1 + |x|^2).$$

Proof. Firstly, we suppose that $s \leq t$, so $Y_s^{t,x}$ is a deterministic variable. We are able to use estimates of the existence Theorem 4.1 and Lemma 4.8:

$$\begin{aligned} -C \log \mathbb{E} \left[\sup_{0 \leq s \leq T} \exp \left(C + \gamma e^{\beta T} (\alpha' + T\alpha) |X_s^{t,x}|^2 \right) \right] &\leq Y_s^{t,x} \leq C \left(1 + \mathbb{E} \left[\sup_{0 \leq s \leq T} |X_s^{t,x}|^4 \right] \right)^{1/2} \\ -\tilde{C}(1 + |x|^2) &\leq Y_s^{t,x} \leq \tilde{C}(1 + |x|^2). \end{aligned}$$

Then, according to the representation theorem, we have

$$\begin{aligned} Y_0^{t,x} &= \mathbb{E}^{\mathbb{Q}^*} \left[h(X_T^{t,x}) + \int_0^T f(s, X_s^{t,x}, Y_s^{t,x}, q_s^*) ds \right] \\ &\geq -C - \alpha' \mathbb{E}^{\mathbb{Q}^*} [|X_T^{t,x}|^2] \\ &\quad + \frac{1}{2\gamma} \mathbb{E}^{\mathbb{Q}^*} \left[\int_0^T |q_u^*|^2 du \right] - \alpha \mathbb{E}^{\mathbb{Q}^*} \left[\int_0^T |X_s^{t,x}|^2 ds \right] - \beta \mathbb{E}^{\mathbb{Q}^*} \left[\int_0^T |Y_s^{t,x}| ds \right]. \end{aligned}$$

But, thanks to the uniqueness, we have $Y_s^{t,x} = Y_s^{s, X_s^{t,x}}$ for $s \geq t$, so $\mathbb{E}^{\mathbb{Q}^*} [|Y_s^{t,x}|] \leq C \left(1 + \mathbb{E}^{\mathbb{Q}^*} [|X_s^{t,x}|^2] \right)$. Moreover, we are allowed to use Lemma 4.11,

$$\begin{aligned} Y_0^{t,x} &\geq -C(1 + |x|^2) - C(\alpha' + T\alpha + \beta C) \left(1 + |x|^2 + T \int_t^T \mathbb{E}^{\mathbb{Q}^*} [|q_u^*|^2] du \right) \\ &\quad + \frac{1}{2\gamma} \mathbb{E}^{\mathbb{Q}^*} \left[\int_0^T |q_u^*|^2 du \right] \\ &\geq -C(1 + |x|^2) + \left(\frac{1}{2\gamma} - CT \right) \mathbb{E}^{\mathbb{Q}^*} \left[\int_0^T |q_u^*|^2 du \right]. \end{aligned}$$

We set $0 < T_2 \leq T_1$ such that $\frac{1}{2\gamma} - CT > 0$ for all $T \in [0, T_2]$. Finally,

$$\mathbb{E}^{\mathbb{Q}^*} \left[\int_0^T |q_u^*|^2 du \right] \leq C(1 + |x|^2 + Y_0^{t,x}) \leq \tilde{C}(1 + |x|^2).$$

□

According to the Proposition 4.14 we know that $\mathbb{E}^{\mathbb{Q}^*} \left[\int_0^T |q_u^*|^2 du \right] \leq \tilde{C}(1 + |x|^2)$ so we are allowed to restrict $\mathcal{A}_{0,T}$: for all $R \geq 0$ we define

$$\mathcal{A}_{0,T}^R = \mathcal{A}_{0,T} \cap \left\{ (q_s)_{s \in [0,T]} : \mathbb{E}^{\mathbb{Q}} \left[\int_0^T |q_u|^2 du \right] \leq \tilde{C}(1 + R^2) \right\}. \quad (4.10)$$

With this new admissible control set we will prove a last inequality:

Proposition 4.15 $\exists C \geq 0, \forall T \in [0, T_2], \forall t, t' \in [0, T], \forall x, x' \in \mathbb{R}^d, \forall q \in \mathcal{A}_{0,T}^{|x| \vee |x'|}, \forall s \in [0, T],$

$$\mathbb{E}^{\mathbb{Q}} \left[\left| X_s^{t,x} - X_s^{t',x'} \right|^2 \right] \leq C \left(|x - x'|^2 + (1 + |x|^2 + |x'|^2) |t - t'| \right).$$

Proof.

$$\mathbb{E}^{\mathbb{Q}} \left[\left| X_s^{t,x} - X_s^{t',x'} \right|^2 \right] \leq 2\mathbb{E}^{\mathbb{Q}} \left[\left| X_s^{t,x} - X_s^{t,x'} \right|^2 \right] + 2\mathbb{E}^{\mathbb{Q}} \left[\left| X_s^{t,x'} - X_s^{t',x'} \right|^2 \right].$$

We have, for $s \geq t$,

$$X_s^{t,x} - X_s^{t,x'} = x - x' + \int_t^s \left(b(u, X_u^{t,x}) - b(u, X_u^{t,x'}) \right) du.$$

So,

$$\mathbb{E}^{\mathbb{Q}} \left[\left| X_s^{t,x} - X_s^{t,x'} \right|^2 \right] \leq C \left(|x - x'|^2 + \int_t^s \mathbb{E}^{\mathbb{Q}} \left[\left| X_u^{t,x} - X_u^{t,x'} \right|^2 \right] du \right).$$

We apply Gronwall's Lemma to obtain that

$$\mathbb{E}^{\mathbb{Q}} \left[\left| X_s^{t,x} - X_s^{t,x'} \right|^2 \right] \leq C |x - x'|^2.$$

Now we deal with the second term. Let us assume that $t \leq t'$. For $s \leq t$, $X_s^{t,x'} - X_s^{t',x'} = 0$. When $t \leq s \leq t'$, we have

$$X_s^{t,x'} - X_s^{t',x'} = \int_t^s b(u, X_u^{t,x'}) du + \int_t^s \sigma(u) dW_u^q + \int_t^s \sigma(u) q_u du.$$

So,

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} \left[\left| X_s^{t,x'} - X_s^{t',x'} \right|^2 \right] &\leq C \left(\mathbb{E}^{\mathbb{Q}} \left[\left(\int_t^{t'} |b(u, X_u^{t,x'})| du \right)^2 \right] + \int_t^{t'} |\sigma(u)|^2 du + \right. \\ &\quad \left. \mathbb{E}^{\mathbb{Q}} \left[\left(\int_t^{t'} |\sigma(u) q_u| du \right)^2 \right] \right) \\ &\leq C \left(|t' - t| + |t' - t| \int_t^{t'} \mathbb{E}^{\mathbb{Q}} \left[\left| X_u^{t,x'} \right|^2 \right] du + |t' - t| \int_t^{t'} \mathbb{E}^{\mathbb{Q}} \left[|q_u|^2 \right] du \right) \\ &\leq C |t' - t| \left(1 + |x'|^2 + \int_0^T \mathbb{E}^{\mathbb{Q}} \left[|q_u|^2 \right] du \right) \\ &\leq C(1 + |x|^2 + |x'|^2) |t' - t|. \end{aligned}$$

Lastly, when $t' \leq s$,

$$X_s^{t,x'} - X_s^{t',x'} = X_{t'}^{t,x'} - X_{t'}^{t',x'} + \int_{t'}^s \left(b(u, X_u^{t,x'}) - b(u, X_u^{t',x'}) \right) du.$$

So,

$$\mathbb{E}^{\mathbb{Q}} \left[\left| X_s^{t,x'} - X_s^{t',x'} \right|^2 \right] \leq C(1 + |x|^2 + |x'|^2) |t' - t| + C \int_{t'}^s \mathbb{E}^{\mathbb{Q}} \left[\left| X_u^{t,x'} - X_u^{t',x'} \right|^2 \right] du,$$

and according to Gronwall's Lemma,

$$\mathbb{E}^{\mathbb{Q}} \left[\left| X_s^{t,x'} - X_s^{t,x} \right|^2 \right] \leq C(1 + |x|^2 + |x'|^2) |t' - t|.$$

□

Proof of Proposition 4.10. First of all, let us assume that $T < T_2$. With this condition, we are allowed to use all previous propositions. Firstly, the quadratic increase of u is already proved in Proposition 4.14. Then, we will show continuity of u in $(t_0, x_0) \in [0, T] \times \mathbb{R}^d$. We have

$$\forall (t, x) \in [0, T] \times \mathbb{R}^d, \quad |u(t, x) - u(t_0, x_0)| \leq |u(t, x) - u(t, x_0)| + |u(t, x_0) - u(t_0, x_0)|.$$

Let us begin with the first term. We define $R := |x| \vee |x_0|$. Thanks to the representation theorem, we have

$$Y_t^{t,x} = \text{ess inf}_{q \in \mathcal{A}_{0,T}^R} Y_t^{q,t,x} \quad \text{and} \quad Y_t^{t,x_0} = \text{ess inf}_{q \in \mathcal{A}_{0,T}^R} Y_t^{q,t,x_0}.$$

So,

$$|Y_t^{t,x} - Y_t^{t,x_0}| \leq \text{ess sup}_{q \in \mathcal{A}_{0,T}^R} |Y_t^{q,t,x} - Y_t^{q,t,x_0}|.$$

But, for $t \leq s \leq T$,

$$\begin{aligned} |Y_s^{q,t,x} - Y_s^{q,t,x_0}| &= \left| \mathbb{E}^{\mathbb{Q}} \left[h(X_T^{t,x}) - h(X_T^{t,x_0}) \right. \right. \\ &\quad \left. \left. + \int_s^T (f(u, X_u^{t,x}, Y_u^{q,t,x}, q_u) - f(u, X_u^{t,x_0}, Y_u^{q,t,x_0}, q_u)) du \middle| \mathcal{F}_s \right] \right|, \\ \mathbb{E}^{\mathbb{Q}} [|Y_s^{q,t,x} - Y_s^{q,t,x_0}|] &\leq \mathbb{E}^{\mathbb{Q}} \left[C(1 + |X_T^{t,x}|^2 + |X_T^{t,x_0}|^2) \right]^{1/2} \mathbb{E}^{\mathbb{Q}} [|X_T^{t,x} - X_T^{t,x_0}|^2]^{1/2} \\ &\quad + \int_s^T \mathbb{E}^{\mathbb{Q}} \left[C(1 + |X_u^{t,x}|^2 + |X_u^{t,x_0}|^2) \right]^{1/2} \mathbb{E}^{\mathbb{Q}} [|X_u^{t,x} - X_u^{t,x_0}|^2]^{1/2} du \\ &\quad + C \int_s^T \mathbb{E}^{\mathbb{Q}} [|Y_u^{q,t,x} - Y_u^{q,t,x_0}|] du, \end{aligned}$$

thanks to Assumption (A.4) and Hölder's inequality. According to Lemma 4.11, the definition of $\mathcal{A}_{0,T}^R$ and Proposition 4.15, we obtain

$$\mathbb{E}^{\mathbb{Q}} [|Y_s^{q,t,x} - Y_s^{q,t,x_0}|] \leq C(1 + |x|^2 + |x_0|^2)^{1/2} |x - x_0| + C \int_s^T \mathbb{E}^{\mathbb{Q}} [|Y_u^{q,t,x} - Y_u^{q,t,x_0}|] du.$$

Then, Gronwall's lemma gives us $|Y_t^{q,t,x} - Y_t^{q,t,x_0}| \leq C(1 + |x| + |x_0|) |x - x_0|$. Since this bound is independent of q , we finally obtain that

$$|Y_t^{t,x} - Y_t^{t,x_0}| \leq C(1 + |x| + |x_0|) |x - x_0|.$$

Now, we will study the second term. Without loss of generality, let us assume that $t < t_0$.

$$\begin{aligned} |Y_t^{t,x_0} - Y_{t_0}^{t_0,x_0}| &\leq |Y_t^{t,x_0} - Y_{t_0}^{t_0,x_0}| + \int_t^{t_0} |g(s, x_0, Y_s^{t_0,x_0}, 0)| ds, \\ &\leq |Y_t^{t,x_0} - Y_{t_0}^{t_0,x_0}| + \int_t^{t_0} C(1 + |x_0|^2 + |Y_s^{t_0,x_0}|) ds. \end{aligned}$$

We apply Proposition 4.14 to obtain

$$|Y_t^{t,x_0} - Y_{t_0}^{t_0,x_0}| \leq |Y_t^{t,x_0} - Y_{t_0}^{t_0,x_0}| + C(1 + |x_0|^2)(t_0 - t).$$

We still have

$$|Y_t^{t,x_0} - Y_{t_0}^{t_0,x_0}| \leq \text{ess sup}_{q \in \mathcal{A}_{0,T}^R} |Y_t^{q,t,x_0} - Y_{t_0}^{q,t_0,x_0}|.$$

Moreover, exactly as the bound estimation for $\mathbb{E}^{\mathbb{Q}} |Y_s^{q,t,x} - Y_s^{q,t,x_0}|$, we have, for $t \leq s \leq T$,

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} [|Y_s^{q,t,x_0} - Y_s^{q,t_0,x_0}|] &\leq \mathbb{E}^{\mathbb{Q}} \left[C(1 + |X_T^{t,x_0}|^2 + |X_T^{t_0,x_0}|^2) \right]^{1/2} \mathbb{E}^{\mathbb{Q}} [|X_T^{t,x_0} - X_T^{t_0,x_0}|^2]^{1/2} \\ &\quad + \int_s^T \mathbb{E}^{\mathbb{Q}} \left[C(1 + |X_u^{t,x_0}|^2 + |X_u^{t_0,x_0}|^2) \right]^{1/2} \mathbb{E}^{\mathbb{Q}} [|X_u^{t,x_0} - X_u^{t_0,x_0}|^2]^{1/2} du \\ &\quad + C \int_s^T \mathbb{E}^{\mathbb{Q}} [|Y_u^{q,t,x_0} - Y_u^{q,t_0,x_0}|] du. \end{aligned}$$

According to Lemma 4.11, the definition of $\mathcal{A}_{0,T}^R$, Proposition 4.15 and Gronwall's Lemma, we obtain $|Y_t^{q,t,x_0} - Y_t^{q,t_0,x_0}| \leq C(1 + |x|^2 + |x_0|^2) |t - t_0|^{1/2}$. Since this bound is independent of q , we finally obtain that

$$|Y_t^{t,x_0} - Y_t^{t_0,x_0}| \leq C(1 + |x|^2 + |x_0|^2) |t - t_0|^{1/2}.$$

So,

$$|u(t, x) - u(t_0, x_0)| \leq C(1 + |x| + |x_0|) |x - x_0| + C(1 + |x|^2 + |x_0|^2) |t - t_0|^{1/2}.$$

We now return to the general case (for T): we set $N \in \mathbb{N}$ such that $T/N < T_2$ and, for $i \in \{0, \dots, N\}$, we define $t_i := iT/N$. According to the beginning of the proof, u is continuous on $[t_{N-1}, T] \times \mathbb{R}^d$. We define $h_{N-1}(x) := Y_{t_{N-1}}^{t_{N-1},x}$. Since $|h_{N-1}(x) - h_{N-1}(x')| \leq C(1 + |x| + |x'|) |x - x'|$, we are allowed to reuse previous results to show the continuity of u on $[t_{N-2}, t_{N-1}] \times \mathbb{R}^d$. Thus, we can iterate this argument to show the continuity of u on $[0, T] \times \mathbb{R}^d$. Moreover the quadratic increase of u with respect to the variable x results from the quadratic increase of u on each interval.

Finally, we will use a stability result to show that u is a viscosity solution to the PDE (4.6). As in the proof of Theorem 4.1, let us consider the function

$$g_n(t, x, y, z) := \inf \{g(t, x, p, q) + n|p - y| + n|q - z|, (p, q) \in \mathbb{Q}^{1+d}\}.$$

We have already seen that $(g_n)_{n \geq [r]}$ is increasing and converges uniformly on compact sets to g . Let $(Y^{n,t,x}, Z^{n,t,x})$ be the unique solution in $\mathcal{S}^2 \times M^2$ to BSDE($h(X_T^{t,x}), -g_n(\cdot, X_\cdot^{t,x}, \cdot, \cdot)$). We define

$$u_n(t, x) := Y_t^{n,t,x}.$$

Then by a classical theorem (see e.g. [73, 41]), u_n is a viscosity solution to the PDE

$$\partial_t u(t, x) + \mathcal{L}u(t, x) - g_n(t, x, u(t, x), -\sigma^* \nabla_x u(t, x)) = 0, \quad u(T, \cdot) = h.$$

Moreover, it follows from the classical comparison theorem that $(u_n)_{n \geq [r]}$ is decreasing and, by construction, converges pointwise to u . Since u is continuous, Dini's theorem implies that the convergence is also uniform on compact sets. Then, we apply a stability result (see e.g. Theorem 1.7. of Chapter 5 in [9]) to prove that u is a viscosity solution to the PDE (4.6). \square

Remark. The uniqueness of viscosity solution to PDE is considered by Da Lio and Ley in [31] and [30].

Acknowledgements. The work of Freddy Delbaen was sponsored by a grant of Credit Suisse as well as by a grant NCCR-Finrisk. The text only reflects the opinion of the authors. The work of Ying Hu was partially supported by funds from the Marie Curie ITN Grant, "Controlled Systems", GA no. 213841/2008. The authors would like to thank an anonymous referee for his/her very careful reading and helpful comments.

Troisième partie

Simulation des EDSRs quadratiques

Chapitre 5

Simulation d'EDSRs dont le générateur est à croissance quadratique

Résumé : Ce chapitre traite de la résolution numérique d'équations différentielles stochastiques rétrogrades markoviennes (EDSRs) dont le générateur à une croissance quadratique par rapport à z et dont la condition terminale est bornée. Nous estimons dans un premier temps des bornes sur le processus Z et nous précisons le théorème de Zhang portant sur la régularité des trajectoires. Nous donnons ensuite un nouveau schéma de discrétisation en temps dont la particularité est que la grille de discrétisation est non uniforme. Enfin nous obtenons une vitesse de convergence pour ce schéma.

Mots clés. Équations différentielles stochastiques rétrogrades, générateur à croissance quadratique, schéma de discrétisation temporelle, grille de discrétisation non uniforme.

Abstract: This chapter deals with the numerical resolution of Markovian backward stochastic differential equations (BSDEs) with drivers of quadratic growth with respect to z and bounded terminal conditions. We first show some bound estimates on the process Z and we specify the Zhang's path regularity theorem. Then we give a new time discretization scheme with a non uniform time net for such BSDEs and we obtain an explicit convergence rate for this scheme.

Key words. Backward stochastic differential equations, driver of quadratic growth, time discretization scheme, non uniform time net.

AMS subject classifications. 60H10, 93E20.

This chapter will be published in
The Annals of Applied Probability
under the title:
Numerical simulation of BSDEs with drivers of quadratic growth

5.1 Introduction

Since the early nineties, there has been an increasing interest for backward stochastic differential equations (BSDEs for short). These equations have a wide range of applications in stochastic control, in finance or in partial differential equation theory. A particular class of BSDE is studied since few years: BSDEs with drivers of quadratic growth with respect to the variable z . This class arises, for example, in the context of utility optimization problems with exponential utility functions, or alternatively in questions related to risk minimization for the entropic risk measure (see e.g. [55]). Many papers deal with existence and uniqueness of solution for such BSDEs: we refer the reader to [60, 64] when the terminal condition is bounded and [24, 25, 36] for the unbounded case. Our concern is rather related to the simulation of BSDEs and more precisely time discretization of BSDEs coupled with a forward stochastic differential equation (SDE for short). Actually, the design of efficient algorithms which are able to solve BSDEs in any reasonable dimension has been intensively studied since the first work of Chevance [28], see for instance [84, 17, 48]. But in all these works, the driver of the BSDE is a Lipschitz function with respect to z and this assumption plays a key role in their proofs. In a recent paper, Cheridito and Stadje [27] studied approximation of BSDEs by backward stochastic difference equations which are based on random walks instead of Brownian motions. They obtain a convergence result when the driver has a subquadratic growth with respect to z and they give an example where this approximation does not converge when the driver has a quadratic growth. To the best of our knowledge, the only work where the time approximation of a BSDE with a quadratic growth with respect to z is studied is the one of Imkeller and Reis [56]. Let notice that, when the driver has a specific form¹, it is possible to get around the problem by using an exponential transformation method (see [57]) or by using results on fully coupled forward-backward differential equations (see [34]).

To explain ideas of this paper, let us introduce (X, Y, Z) the solution to the forward backward system

$$\begin{aligned} X_t &= x + \int_0^t b(s, X_s)ds + \int_0^t \sigma(s)dW_s, \\ Y_t &= g(X_T) + \int_t^T f(s, X_s, Y_s, Z_s)ds - \int_t^T Z_s dW_s, \end{aligned}$$

where g is bounded, f is locally Lipschitz and has a quadratic growth with respect to z . A well-known result is that when g is a Lipschitz function with Lipschitz constant K_g , then the process Z is bounded by $C(K_g + 1)$ (see Theorem 5.3). So, in this case, the driver of the BSDE is a Lipschitz function with respect to z . Thereby, a simple idea is to do an approximation of (Y, Z) by the solution (Y^N, Z^N) to the BSDE

$$Y_t^N = g_N(X_T) + \int_t^T f(s, X_s, Y_s^N, Z_s^N)ds - \int_t^T Z_s^N dW_s,$$

where g_N is a Lipschitz approximation of g . Thanks to bounded mean oscillation martingale (BMO martingale in the sequel) tools, we have an error estimate for this approximation: see e.g. [56, 20] or Proposition 5.9. For example, if g is α -Hölder, we are able to obtain the error bound $CK_{g_N}^{\frac{-\alpha}{1-\alpha}}$ (see Proposition 5.17). Moreover, we can have an error estimate for the time discretization of the approximated BSDE thanks to any numerical scheme for BSDEs with Lipschitz driver. But, this error estimate depends on K_{g_N} : roughly speaking, this error is $Ce^{CK_{g_N}^2 n^{-1}}$ with n the number of discretization times. The exponential term results from the use of Gronwall's inequality. Finally, when g is α -Hölder and $K_{g_N} = N$, the global error bound is

$$C \left(\frac{1}{N^{\frac{\alpha}{1-\alpha}}} + \frac{e^{CN^2}}{n} \right). \quad (5.1)$$

So, when N increases, n^{-1} will have to become small very quickly and the speed of convergence turns out to be bad: if we take $N = (\frac{C}{\varepsilon} \log n)^{1/2}$ with $0 < \varepsilon < 1$, then the global error bound becomes $C_\varepsilon (\log n)^{\frac{-\alpha}{2(1-\alpha)}}$. The same drawback appears in the work of Imkeller and Reis [56]. Indeed, their idea is to do an approximation of (Y, Z) by the solution (Y^N, Z^N) to the truncated BSDE

$$Y_t^N = g(X_T) + \int_t^T f(s, X_s, Y_s^N, h_N(Z_s^N))ds - \int_t^T Z_s^N dW_s,$$

¹Roughly speaking, the driver is a sum of a quadratic term $z \mapsto C|z|^2$ and a function that has a linear growth with respect to z .

where $h_N : \mathbb{R}^{1 \times d} \rightarrow \mathbb{R}^{1 \times d}$ is a smooth modification of the projection on the open Euclidean ball of radius N about 0. Thanks to several statements concerning the path regularity and stochastic smoothness of the solution processes, the authors show that for any $\beta \geq 1$, the approximation error is lower than $C_\beta N^{-\beta}$. So, they obtain the global error bound

$$C_\beta \left(\frac{1}{N^\beta} + \frac{e^{CN^2}}{n} \right), \quad (5.2)$$

and, consequently, the speed of convergence also turns out to be bad: if we take $N = (\frac{C}{\varepsilon} \log n)^{1/2}$ with $0 < \varepsilon < 1$, then the global error bound becomes $C_{\beta, \varepsilon} (\log n)^{-\beta/2}$.

Another idea is to use an estimate of Z that does not depends on K_g . So, we extend a result of [35] which shows

$$|Z_t| \leq M_1 + \frac{M_2}{(T-t)^{1/2}}, \quad 0 \leq t < T. \quad (5.3)$$

Let us notice that this type of estimation is well known in the case of drivers with linear growth as a consequence of the Bismut-Elworthy formula: see e.g. [44]. But in our case, we do not need to suppose that σ is invertible. Then, thanks to this estimation, we know that, when $t < T$, $f(t, \cdot, \cdot, \cdot)$ is a Lipschitz function with respect to z and the Lipschitz constant depends on t . So we are able to modify the classical uniform time net to obtain a convergence speed for a modified time discretization scheme for our BSDE: the idea is to put more discretization points near the final time T than near 0. The same idea is used by Gobet and Makhlof in [51] for BSDEs with drivers of linear growth and a terminal function g not Lipschitz. But due to technical reasons we need to apply this modified time discretization scheme to the approximated BSDE:

$$Y_t^{N, \varepsilon} = g_N(X_T) + \int_t^T f^\varepsilon(s, X_s, Y_s^{N, \varepsilon}, Z_s^{N, \varepsilon}) ds - \int_t^T Z_s^{N, \varepsilon} dW_s,$$

with

$$f^\varepsilon(s, x, y, z) := \mathbb{1}_{s \leq T-\varepsilon} f(s, x, y, z) + \mathbb{1}_{s > T-\varepsilon} f(s, x, y, 0).$$

Thanks to the estimate (5.3), we obtain a speed convergence for the time discretization scheme of this approximated BSDE (see Theorem 5.15). Moreover, BMO tools give us again an estimate of the approximation error (see Proposition 5.9). Finally, if we suppose that g is α -Hölder, we prove that we can choose properly N and ε to obtain the global error estimate $C n^{-\frac{2\alpha}{(2-\alpha)(2+K)-2+2\alpha}}$ (see Theorem 5.20) where $K > 0$ depends on constant M_2 defined in equation (5.3) and constants related to f . Let us notice that such a speed of convergence where constants related to f , g , b and σ appear in the power of n is unusual. Even if we have an error far better than (5.1) or (5.2), this result is not very interesting in practice because the speed of convergence strongly depends on K . But, when b is bounded, we prove that we can take M_2 as small as we want in (5.3). Finally, we obtain a global error estimate lower than $C_\eta n^{-(\alpha-\eta)}$, for all $\eta > 0$ (see Theorem 5.23).

The paper is organized as follows. In the introductory Section 2 we recall some of the well known results concerning SDEs and BSDEs. In Section 3 we establish some estimates concerning the process Z : we show a first uniform bound for Z , then a time dependent bound and finally we specify the classical path regularity theorem. In Section 4 we define a modified time discretization scheme for BSDEs with a non uniform time net and we obtain an explicit error bound.

5.2 Preliminaries

5.2.1 Notations

Throughout this paper, $(W_t)_{t \geq 0}$ will denote a d -dimensional Brownian motion, defined on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. For $t \geq 0$, let \mathcal{F}_t denote the σ -algebra $\sigma(W_s; 0 \leq s \leq t)$, augmented with the \mathbb{P} -null sets of \mathcal{F} . The Euclidian norm on \mathbb{R}^d will be denoted by $|\cdot|$. The operator norm induced by $|\cdot|$ on the space of linear operator is also denoted by $|\cdot|$. For $p \geq 2$, $m \in \mathbb{N}$, we denote further

- $\mathcal{S}^p(\mathbb{R}^m)$, or \mathcal{S}^p when no confusion is possible, the space of all adapted processes $(Y_t)_{t \in [0, T]}$ with values in \mathbb{R}^m normed by $\|Y\|_{\mathcal{S}^p} = \mathbb{E}[(\sup_{t \in [0, T]} |Y_t|)^p]^{1/p}$; $\mathcal{S}^\infty(\mathbb{R}^m)$, or \mathcal{S}^∞ , the space of bounded measurable processes;

- $\mathcal{M}^p(\mathbb{R}^m)$, or \mathcal{M}^p , the space of all progressively measurable processes $(Z_t)_{t \in [0, T]}$ with values in \mathbb{R}^m normed by $\|Z\|_{\mathcal{M}^p} = \mathbb{E}[(\int_0^T |Z_s|^2 ds)^{p/2}]^{1/p}$.

In the following, we keep the same notation C for all finite, nonnegative constants that appear in our computations: they may depend on known parameters deriving from assumptions and on T , but not on any of the approximation and discretization parameters. In the same spirit, we keep the same notation η for all finite, positive constants that we can take as small as we want independently of the approximation and discretization parameters.

5.2.2 Some results on BMO martingales

In our work, the space of BMO martingales play a key role for the a priori estimates needed in our analysis of BSDEs. We refer the reader to [58] for the theory of BMO martingales and we just recall the properties that we will use in the sequel. Let $\Phi_t = \int_0^t \phi_s dW_s$, $t \in [0, T]$ be a real square integrable martingale with respect to the Brownian filtration. Then Φ is a BMO martingale if

$$\|\Phi\|_{BMO} = \sup_{\tau \in [0, T]} \mathbb{E}[\langle \Phi \rangle_T - \langle \Phi \rangle_\tau | \mathcal{F}_\tau]^{1/2} = \sup_{\tau \in [0, T]} \mathbb{E} \left[\int_\tau^T \phi_s^2 ds | \mathcal{F}_\tau \right]^{1/2} < +\infty,$$

where the supremum is taken over all stopping times in $[0, T]$; $\langle \Phi \rangle$ denotes the quadratic variation of Φ . In our case, the very important feature of BMO martingales is the following lemma:

Lemma 5.1 *Let Φ be a BMO martingale. Then we have:*

1. *The stochastic exponential*

$$\mathcal{E}(\Phi)_t = \mathcal{E}_t = \exp \left(\int_0^t \phi_s dW_s - \frac{1}{2} \int_0^t |\phi_s|^2 ds \right), \quad 0 \leq t \leq T,$$

is a uniformly integrable martingale.

2. *Thanks to the reverse Hölder inequality, there exists $p > 1$ such that $\mathcal{E}_T \in L^p$. The maximal p with this property can be expressed in terms of the BMO norm of Φ .*

$$3. \quad \forall n \in \mathbb{N}^*, \mathbb{E} \left[\left(\int_0^T |\phi_s|^2 ds \right)^n \right] \leq n! \|\Phi\|_{BMO}^{2n}.$$

5.2.3 The backward-forward system

Given functions b, σ, g and f , for $x \in \mathbb{R}^d$ we will deal with the solution (X, Y, Z) to the following system of (decoupled) backward-forward stochastic differential equations: for $t \in [0, T]$,

$$X_t = x + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s) dW_s, \quad (5.4)$$

$$Y_t = g(X_T) + \int_t^T f(s, X_s, Y_s, Z_s) ds - \int_t^T Z_s dW_s. \quad (5.5)$$

For the functions that appear in the above system of equations we give some general assumptions.

(HX0). $b : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$, $\sigma : [0, T] \rightarrow \mathbb{R}^{d \times d}$ are measurable functions. There exist four positive constants M_b, K_b, M_σ and K_σ such that $\forall t, t' \in [0, T], \forall x, x' \in \mathbb{R}^d$,

$$\begin{aligned} |b(t, x)| &\leq M_b(1 + |x|), \\ |b(t, x) - b(t', x')| &\leq K_b(|x - x'| + |t - t'|^{1/2}), \\ |\sigma(t)| &\leq M_\sigma, \\ |\sigma(t) - \sigma(t')| &\leq K_\sigma |t - t'|. \end{aligned}$$

(HY0). $f : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^{1 \times d} \rightarrow \mathbb{R}$, $g : \mathbb{R}^d \rightarrow \mathbb{R}$ are measurable functions. There exist five positive constants M_f , $K_{f,x}$, $K_{f,y}$, $K_{f,z}$ and M_g such that $\forall t \in [0, T]$, $\forall x, x' \in \mathbb{R}^d$, $\forall y, y' \in \mathbb{R}$, $\forall z, z' \in \mathbb{R}^{1 \times d}$,

$$\begin{aligned} |f(t, x, y, z)| &\leq M_f(1 + |y| + |z|^2), \\ |f(t, x, y, z) - f(t, x', y', z')| &\leq K_{f,x}|x - x'| + K_{f,y}|y - y'| + (K_{f,z} + L_{f,z}(|z| + |z'|))|z - z'|, \\ |g(x)| &\leq M_g. \end{aligned}$$

We next recall some results on BSDEs with quadratic growth. For their original version and their proof we refer to [60], [20] and [56].

Theorem 5.2 *Under (HX0), (HY0), the system (5.4)-(5.5) has a unique solution $(X, Y, Z) \in \mathcal{S}^2 \times \mathcal{S}^\infty \times \mathcal{M}^2$. The martingale $Z * W$ belongs to the space of BMO martingales and $\|Z * W\|_{BMO}$ only depends on T , M_g and M_f . Moreover, there exists $r > 1$ such that $\mathcal{E}(Z * W) \in L^r$.*

5.3 Some useful estimates of Z

5.3.1 A first bound for Z

Theorem 5.3 *Suppose that (HX0), (HY0) hold and that g is Lipschitz with Lipschitz constant K_g . Then, there exists a version of Z such that, $\forall t \in [0, T]$,*

$$|Z_t| \leq e^{(2K_b + K_{f,y})T} M_\sigma(K_g + TK_{f,x}).$$

Proof. Firstly, we suppose that b , g and f are differentiable with respect to x , y and z . Then (X, Y, Z) is differentiable with respect to x and $(\nabla X, \nabla Y, \nabla Z)$ is solution of

$$\nabla X_t = I_d + \int_0^t \nabla b(s, X_s) \nabla X_s ds, \quad (5.6)$$

$$\begin{aligned} \nabla Y_t &= \nabla g(X_T) \nabla X_T - \int_t^T \nabla Z_s dW_s \\ &\quad + \int_t^T \nabla_x f(s, X_s, Y_s, Z_s) \nabla X_s + \nabla_y f(s, X_s, Y_s, Z_s) \nabla Y_s + \nabla_z f(s, X_s, Y_s, Z_s) \nabla Z_s ds, \end{aligned} \quad (5.7)$$

where $\nabla X_t = (\partial X_t^i / \partial x^j)_{1 \leq i, j \leq d}$, $\nabla Y_t = {}^t(\partial Y_t / \partial x^j)_{1 \leq j \leq d} \in \mathbb{R}^{1 \times d}$, $\nabla Z_t = (\partial Z_t^i / \partial x^j)_{1 \leq i, j \leq d}$ and $\int_t^T \nabla Z_s dW_s$ means

$$\sum_{1 \leq i \leq d} \int_t^T (\nabla Z_s)^i dW_s^i$$

with $(\nabla Z)^i$ denoting the i -th line of the $d \times d$ matrix process ∇Z . Thanks to usual transformations on the BSDE we obtain

$$\begin{aligned} e^{\int_0^t \nabla_y f(s, X_s, Y_s, Z_s) ds} \nabla Y_t &= e^{\int_0^T \nabla_y f(s, X_s, Y_s, Z_s) ds} \nabla g(X_T) \nabla X_T - \int_t^T e^{\int_0^s \nabla_y f(u, X_u, Y_u, Z_u) du} \nabla Z_s d\tilde{W}_s \\ &\quad + \int_t^T e^{\int_0^s \nabla_y f(u, X_u, Y_u, Z_u) du} \nabla_x f(s, X_s, Y_s, Z_s) \nabla X_s ds, \end{aligned}$$

with $d\tilde{W}_s = dW_s - \nabla_z f(s, X_s, Y_s, Z_s) ds$. We have

$$\begin{aligned} \left\| \int_0^T \nabla_z f(s, X_s, Y_s, Z_s) dW_s \right\|_{BMO}^2 &= \sup_{\tau \in [0, T]} \mathbb{E} \left[\int_\tau^T |\nabla_z f(s, X_s, Y_s, Z_s)|^2 ds \middle| \mathcal{F}_\tau \right] \\ &\leq C \left(1 + \sup_{\tau \in [0, T]} \mathbb{E} \left[\int_\tau^T |Z_s|^2 ds \middle| \mathcal{F}_\tau \right] \right) \\ &= C \left(1 + \|Z * W\|_{BMO}^2 \right). \end{aligned}$$

Since $Z * W$ belongs to the space of BMO martingales, $\|\int_0^\cdot \nabla_z f(s, X_s, Y_s, Z_s) dW_s\|_{BMO} < +\infty$. Lemma 5.1 gives us that $\mathcal{E}(\int_0^\cdot \nabla_z f(s, X_s, Y_s, Z_s) dW_s)_t$ is a uniformly integrable martingale, so we are able to apply Girsanov's theorem: there exists a probability \mathbb{Q} under which $(\tilde{W})_{t \in [0, T]}$ is a Brownian motion. Then,

$$\begin{aligned} e^{\int_0^t \nabla_y f(s, X_s, Y_s, Z_s) ds} \nabla Y_t &= \mathbb{E}^{\mathbb{Q}} \left[e^{\int_0^T \nabla_y f(s, X_s, Y_s, Z_s) ds} \nabla g(X_T) \nabla X_T \right. \\ &\quad \left. + \int_t^T e^{\int_0^s \nabla_y f(u, X_u, Y_u, Z_u) du} \nabla_x f(s, X_s, Y_s, Z_s) \nabla X_s ds \middle| \mathcal{F}_t \right], \end{aligned}$$

and

$$|\nabla Y_t| \leq e^{(K_b + K_{f,y})T} (K_g + T K_{f,x}), \quad (5.8)$$

because $|\nabla X_t| \leq e^{K_b T}$. Moreover, thanks to the Malliavin calculus, it is classical to show that a version of $(Z_t)_{t \in [0, T]}$ is given by $(\nabla Y_t (\nabla X_t)^{-1} \sigma(t))_{t \in [0, T]}$. So we obtain

$$|Z_t| \leq e^{K_b T} M_\sigma |\nabla Y_t| \leq e^{(2K_b + K_{f,y})T} M_\sigma (K_g + T K_{f,x}), \quad a.s.,$$

because $|\nabla X_t^{-1}| \leq e^{K_b T}$.

When b, g and f are not differentiable, we can also prove the result by a standard approximation and stability results for BSDEs with linear growth. \square

5.3.2 A time dependent estimate of Z

We will introduce two alternative assumptions.

(HX1). b is differentiable with respect to x and σ is differentiable with respect to t . There exists $\lambda \in \mathbb{R}^+$ such that $\forall \eta \in \mathbb{R}^d$

$$\left| {}^t \eta \sigma(s) [{}^t \sigma(s) {}^t \nabla b(s, x) - {}^t \sigma'(s)] \eta \right| \leq \lambda |{}^t \eta \sigma(s)|^2. \quad (5.9)$$

(HX1'). σ is invertible and $\forall t \in [0, T], |\sigma(t)^{-1}| \leq M_{\sigma^{-1}}$.

Example. Assumption (HX1) is verified when, $\forall s \in [0, T], \nabla b(s, \cdot)$ commutes with $\sigma(s)$ and $\exists A : [0, T] \rightarrow \mathbb{R}^{d \times d}$ bounded such that $\sigma'(t) = \sigma(t) A(t)$.

Theorem 5.4 Suppose that (HX0), (HY0) hold and that (HX1) or (HX1') holds. Moreover, suppose that g is lower (or upper) semi-continuous. Then there exists a version of Z and there exist two constants $C, C' \in \mathbb{R}^+$ that depend only in $T, M_g, M_f, K_{f,x}, K_{f,y}, K_{f,z}$ and $L_{f,z}$ such that, $\forall t \in [0, T]$,

$$|Z_t| \leq C + C'(T - t)^{-1/2}.$$

Proof. In a first time, we will suppose that (HX1) holds and that f, g are differentiable with respect to x, y and z . Then (Y, Z) is differentiable with respect to x and $(\nabla Y, \nabla Z)$ is the solution of the BSDE

$$\begin{aligned} \nabla Y_t &= \nabla g(X_T) \nabla X_T - \int_t^T \nabla Z_s dW_s \\ &\quad + \int_t^T \nabla_x f(s, X_s, Y_s, Z_s) \nabla X_s + \nabla_y f(s, X_s, Y_s, Z_s) \nabla Y_s + \nabla_z f(s, X_s, Y_s, Z_s) \nabla Z_s ds. \end{aligned}$$

Thanks to usual transformations we obtain

$$\begin{aligned} e^{\int_0^t \nabla_y f(s, X_s, Y_s, Z_s) ds} \nabla Y_t + \int_0^t e^{\int_0^s \nabla_y f(u, X_u, Y_u, Z_u) du} \nabla_x f(s, X_s, Y_s, Z_s) \nabla X_s ds &= \\ e^{\int_0^T \nabla_y f(s, X_s, Y_s, Z_s) ds} \nabla g(X_T) \nabla X_T + \int_0^T e^{\int_0^s \nabla_y f(u, X_u, Y_u, Z_u) du} \nabla_x f(s, X_s, Y_s, Z_s) \nabla X_s ds & \\ - \int_t^T e^{\int_0^s \nabla_y f(u, X_u, Y_u, Z_u) du} \nabla Z_s d\tilde{W}_s, & \end{aligned}$$

with $d\tilde{W}_s = dW_s - \nabla_z f(s, X_s, Y_s, Z_s)ds$. We can rewrite it as

$$F_t = F_T - \int_t^T e^{\int_0^s \nabla_y f(u, X_u, Y_u, Z_u)du} \nabla Z_s d\tilde{W}_s \quad (5.10)$$

with

$$F_t := e^{\int_0^t \nabla_y f(s, X_s, Y_s, Z_s)ds} \nabla Y_t + \int_0^t e^{\int_0^s \nabla_y f(u, X_u, Y_u, Z_u)du} \nabla_x f(s, X_s, Y_s, Z_s) \nabla X_s ds.$$

$Z * W$ belongs to the space of BMO martingales so we are able to apply Girsanov's theorem: there exists a probability \mathbb{Q} under which $(\tilde{W})_{t \in [0, T]}$ is a Brownian motion. Thanks to the Malliavin calculus, it is possible to show that $(\nabla Y_t (\nabla X_t)^{-1} \sigma(t))_{t \in [0, T]}$ is a version of Z . Now we define:

$$\begin{aligned} \alpha_t &:= \int_0^t e^{\int_0^s \nabla_y f(u, X_u, Y_u, Z_u)du} \nabla_x f(s, X_s, Y_s, Z_s) \nabla X_s ds (\nabla X_t)^{-1} \sigma(t), \\ \tilde{Z}_t &:= F_t (\nabla X_t)^{-1} \sigma(t) = e^{\int_0^t \nabla_y f(s, X_s, Y_s, Z_s)ds} Z_t + \alpha_t, \quad a.s., \\ \tilde{F}_t &:= e^{\lambda t} F_t (\nabla X_t)^{-1}. \end{aligned}$$

Since $d\nabla X_t = \nabla b(t, X_t) \nabla X_t dt$, then $d(\nabla X_t)^{-1} = -(\nabla X_t)^{-1} \nabla b(t, X_t) dt$ and thanks to Itô's formula,

$$d\tilde{Z}_t = dF_t (\nabla X_t)^{-1} \sigma(t) - F_t (\nabla X_t)^{-1} \nabla b(t, X_t) \sigma(t) dt + F_t (\nabla X_t)^{-1} \sigma'(t) dt,$$

and

$$d(e^{\lambda t} \tilde{Z}_t) = \tilde{F}_t (\lambda Id - \nabla b(t, X_t)) \sigma(t) dt + \tilde{F}_t \sigma'(t) dt + e^{\lambda t} dF_t (\nabla X_t)^{-1} \sigma(t).$$

Finally,

$$d \left| e^{\lambda t} \tilde{Z}_t \right|^2 = d\langle M \rangle_t + 2 \left[\lambda \left| \tilde{F}_t \sigma(t) \right|^2 - \tilde{F}_t \sigma(t) [{}^t \sigma(t)^t \nabla b(t, X_t) - {}^t \sigma'(t)] {}^t \tilde{F}_t \right] dt + dM_t^*,$$

with $M_t := \int_0^t e^{\lambda s} dF_s (\nabla X_s)^{-1} \sigma(s)$ and M_t^* a \mathbb{Q} -martingale. Thanks to the assumption (HX1) we are able to conclude that $\left| e^{\lambda t} \tilde{Z}_t \right|^2$ is a \mathbb{Q} -submartingale. Hence,

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} \left[\int_t^T e^{2\lambda s} \left| \tilde{Z}_s \right|^2 ds \middle| \mathcal{F}_t \right] &\geq e^{2\lambda t} \left| \tilde{Z}_t \right|^2 (T - t) \\ &\geq e^{2\lambda t} \left| e^{\int_0^t \nabla_y f(s, X_s, Y_s, Z_s)ds} Z_t + \alpha_t \right|^2 (T - t) \quad a.s., \end{aligned}$$

which implies

$$\begin{aligned} |Z_t|^2 (T - t) &= e^{-2\lambda t} e^{-2 \int_0^t \nabla_y f(s, X_s, Y_s, Z_s)ds} e^{2\lambda t} \left| e^{\int_0^t \nabla_y f(s, X_s, Y_s, Z_s)ds} Z_t + \alpha_t - \alpha_t \right|^2 (T - t) \\ &\leq C \left(e^{2\lambda t} \left| e^{\int_0^t \nabla_y f(s, X_s, Y_s, Z_s)ds} Z_t + \alpha_t \right|^2 + 1 \right) (T - t) \\ &\leq C \left(\mathbb{E}^{\mathbb{Q}} \left[\int_t^T e^{2\lambda s} \left| \tilde{Z}_s \right|^2 ds \middle| \mathcal{F}_t \right] + (T - t) \right) \quad a.s., \end{aligned}$$

with C a constant that only depends on $T, K_b, M_\sigma, K_{f,x}, K_{f,y}$ and λ . Moreover, we have, a.s.,

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} \left[\int_t^T e^{2\lambda s} \left| \tilde{Z}_s \right|^2 ds \middle| \mathcal{F}_t \right] &\leq C \mathbb{E}^{\mathbb{Q}} \left[\int_t^T |Z_s|^2 + |\alpha_s|^2 ds \middle| \mathcal{F}_t \right] \\ &\leq C \left(\|Z\|_{BMO(\mathbb{Q})}^2 + (T - t) \right). \end{aligned}$$

But $\|Z\|_{BMO(\mathbb{Q})}$ does not depend on K_g because (Y, Z) is a solution of the following quadratic BSDE:

$$Y_t = g(X_T) + \int_t^T (f(s, X_s, Y_s, Z_s) - Z_s \nabla_z f(s, X_s, Y_s, Z_s)) ds - \int_t^T Z_s d\tilde{W}_s. \quad (5.11)$$

Finally $|Z_t| \leq C(1 + (T - t)^{-1/2})$ a.s..

When σ is invertible, the inequality (5.9) is verified with $\lambda := M_{\sigma^{-1}}(M_{\sigma}K_b + K_{\sigma})$. Since this λ does not depend on ∇b and σ' , we can prove the result when $b(t, \cdot)$ and σ are not differentiable by a standard approximation and stability results for BSDEs with linear growth. So, we are allowed to replace assumption (HX1) by (HX1').

When f is not differentiable and g is only Lipschitz we can prove the result by a standard approximation and stability results for linear BSDEs. But we notice that our estimation on Z does not depend on K_g . This allows us to weaken the hypothesis on g further: when g is only lower or upper semi-continuous the result stays true. The proof is the same as the proof of Proposition 4.3 in [35]. \square

Remark 5.5 *The previous proof gives us a more precise estimation for a version of Z when f is differentiable with respect to z : $\forall t \in [0, T]$,*

$$|Z_t| \leq C + C' \mathbb{E}^{\mathbb{Q}} \left[\int_t^T |Z_s|^2 ds \middle| \mathcal{F}_t \right]^{1/2} (T - t)^{-1/2}.$$

Remark 5.6 *When assumptions (HX1) or (HX1') are not verified, the process Z may blow up before T . Zhang gives an example of such a phenomenon in dimension 1: we refer the reader to example 1 in [85].*

5.3.3 Zhang's path regularity Theorem

Let $0 = t_0 < t_1 < \dots < t_n = T$ be any given partition of $[0, T]$, and denote δ_n the mesh size of this partition. We define a set of random variables

$$\bar{Z}_{t_i} = \frac{1}{t_{i+1} - t_i} \mathbb{E} \left[\int_{t_i}^{t_{i+1}} Z_s ds \middle| \mathcal{F}_{t_i} \right], \quad \forall i \in \{0, \dots, n-1\}.$$

Then we are able to precise Theorem 3.4.3 in [86]:

Theorem 5.7 *Suppose that (HX0), (HY0) hold and g is a Lipschitz function, with Lipschitz constant K_g . Then we have*

$$\sum_{i=0}^{n-1} \mathbb{E} \left[\int_{t_i}^{t_{i+1}} |Z_t - \bar{Z}_{t_i}|^2 dt \right] \leq C(1 + K_g^2) \delta_n,$$

where C is a positive constant independent of δ_n and K_g .

Proof. We will follow the proof of Theorem 5.6., in [56]: we just need to specify how the estimate depends on K_g . Firstly, it is not difficult to show that \bar{Z}_{t_i} is the best \mathcal{F}_{t_i} -measurable approximation of Z in $\mathcal{M}^2([t_i, t_{i+1}])$, i.e.

$$\mathbb{E} \left[\int_{t_i}^{t_{i+1}} |Z_t - \bar{Z}_{t_i}|^2 dt \right] = \inf_{Z_i \in L^2(\Omega, \mathcal{F}_{t_i})} \mathbb{E} \left[\int_{t_i}^{t_{i+1}} |Z_t - Z_i|^2 dt \right].$$

In particular,

$$\mathbb{E} \left[\int_{t_i}^{t_{i+1}} |Z_t - \bar{Z}_{t_i}|^2 dt \right] \leq \mathbb{E} \left[\int_{t_i}^{t_{i+1}} |Z_t - Z_{t_i}|^2 dt \right].$$

In the same spirit as previous proofs, we suppose in a first time that b, g and f are differentiable with respect to x, y and z . So,

$$Z_t - Z_{t_i} = \nabla Y_t (\nabla X_t)^{-1} \sigma(t) - \nabla Y_{t_i} (\nabla X_{t_i})^{-1} \sigma(t_i) = I_1 + I_2 + I_3, \quad a.s.,$$

with $I_1 = \nabla Y_t (\nabla X_t)^{-1} (\sigma(t) - \sigma(t_i))$, $I_2 = \nabla Y_t ((\nabla X_t)^{-1} - (\nabla X_{t_i})^{-1}) \sigma(t_i)$ and $I_3 = \nabla (Y_t - Y_{t_i}) (\nabla X_{t_i})^{-1} \sigma(t_i)$. Firstly, thanks to the estimation (5.8) we have

$$|I_1|^2 \leq |\nabla Y_t|^2 e^{2K_b T} K_{\sigma}^2 |t_{i+1} - t_i|^2 \leq C(1 + K_g^2) \delta_n^2.$$

We obtain the same estimation for $|I_2|$ because

$$|(\nabla X_t)^{-1} - (\nabla X_{t_i})^{-1}| \leq \left| \int_{t_i}^t (\nabla X_s)^{-1} \nabla b(s, X_s) ds \right| \leq K_b e^{K_b T} |t - t_i|.$$

Lastly, $|I_3| \leq M_\sigma e^{K_b T} |\nabla Y_t - \nabla Y_{t_i}|$. So,

$$\sum_{i=0}^{n-1} \mathbb{E} \left[\int_{t_i}^{t_{i+1}} |I_3|^2 dt \right] \leq C \delta_n \sum_{i=0}^{n-1} \mathbb{E} \left[\text{ess sup}_{t \in [t_i, t_{i+1}]} |\nabla Y_t - \nabla Y_{t_i}|^2 \right].$$

By using the BSDE (5.7), (HY0), the estimate on ∇X_s and the estimate (5.8), we have

$$\begin{aligned} & |\nabla Y_t - \nabla Y_{t_i}|^2 \\ & \leq C \left(\int_{t_i}^t (C(1 + K_g) + |\nabla_z f(s, X_s, Y_s, Z_s)| |\nabla Z_s|) ds \right)^2 + C \left(\int_{t_i}^t \nabla Z_s dW_s \right)^2. \end{aligned}$$

The inequalities of Hölder and Burkholder-Davis-Gundy give us

$$\begin{aligned} & \sum_{i=0}^{n-1} \mathbb{E} \left[\text{ess sup}_{t \in [t_i, t_{i+1}]} |\nabla Y_t - \nabla Y_{t_i}|^2 \right] \\ & \leq C(1 + K_g^2) + C \sum_{i=0}^{n-1} \mathbb{E} \left(\int_{t_i}^{t_{i+1}} |\nabla_z f(s, X_s, Y_s, Z_s)| |\nabla Z_s| ds \right)^2 + C \mathbb{E} \left(\int_{t_i}^{t_{i+1}} |\nabla Z_s|^2 ds \right) \\ & \leq C(1 + K_g^2) + C \mathbb{E} \left[\left(\int_0^T |\nabla_z f(s, X_s, Y_s, Z_s)| |\nabla Z_s| ds \right)^2 + \int_0^T |\nabla Z_s|^2 ds \right] \\ & \leq C(1 + K_g^2) + C \mathbb{E} \left[\left(\int_0^T (1 + |Z_s|^2) ds \right) \left(\int_0^T |\nabla Z_s|^2 ds \right) + \int_0^T |\nabla Z_s|^2 ds \right] \\ & \leq C(1 + K_g^2) + C \left(1 + \mathbb{E} \left[\left(\int_0^T |Z_s|^2 ds \right)^{p-1/p} \right] \right) \mathbb{E} \left[\left(\int_0^T |\nabla Z_s|^2 ds \right)^q \right]^{1/q}, \end{aligned}$$

for all $p > 1$ and $q > 1$ such that $1/p + 1/q = 1$. But, $(\nabla Y, \nabla Z)$ is solution of BSDE (5.7), so, from Corollary 9 in [20], there exists q that only depends on $\|Z * W\|_{BMO}$ such that

$$\mathbb{E} \left[\left(\int_0^T |\nabla Z_s|^2 ds \right)^q \right]^{1/q} \leq C(1 + K_g^2).$$

Moreover, we can apply Lemma 5.1 to obtain the estimate

$$\mathbb{E} \left[\left(\int_0^T |Z_s|^2 ds \right)^p \right]^{1/p} \leq C \|Z\|_{BMO}^2 \leq C.$$

Finally,

$$\sum_{i=0}^{n-1} \mathbb{E} \left[\int_{t_i}^{t_{i+1}} |I_3|^2 dt \right] \leq C(1 + K_g^2) \delta_n$$

and

$$\begin{aligned} \sum_{i=0}^{n-1} \mathbb{E} \left[\int_{t_i}^{t_{i+1}} |Z_t - \bar{Z}_{t_i}|^2 dt \right] & \leq \sum_{i=0}^{n-1} \mathbb{E} \left[\int_{t_i}^{t_{i+1}} (|I_1|^2 + |I_2|^2 + |I_3|^2) dt \right] \\ & \leq C(1 + K_g^2) \delta_n. \end{aligned}$$

□

5.4 Convergence of a modified time discretization scheme for the BSDE

5.4.1 An approximation of the quadratic BSDE

In a first time we will approximate our quadratic BSDE (5.5) by another one. We set $\varepsilon \in]0, T[$ and $N \in \mathbb{N}$. Let $(Y_t^{N,\varepsilon}, Z_t^{N,\varepsilon})$ the solution of the BSDE

$$Y_t^{N,\varepsilon} = g_N(X_T) + \int_t^T f^\varepsilon(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon}) ds - \int_t^T Z_s^{N,\varepsilon} dW_s, \quad (5.12)$$

with

$$f^\varepsilon(s, x, y, z) := \mathbb{1}_{s \leq T-\varepsilon} f(s, x, y, z) + \mathbb{1}_{s > T-\varepsilon} f(s, x, y, 0),$$

and g_N a Lipschitz approximation of g with Lipschitz constant N . f^ε verifies assumption (HY0) with the same constants as f . Since g_N is a Lipschitz function, $Z^{N,\varepsilon}$ has a bounded version and the BSDE (5.12) is a BSDE with a linear growth. Moreover, we can apply Theorem 5.4 to obtain:

Proposition 5.8 *Let us assume that (HX0), (HY0) and (HX1) or (HX1') hold. There exists a version of $Z^{N,\varepsilon}$ and there exist three constants $M_{z,1}, M_{z,2}, M_{z,3} \in \mathbb{R}^+$ that do not depend on N and ε such that, $\forall s \in [0, T]$,*

$$|Z_s^{N,\varepsilon}| \leq \left(M_{z,1} + \frac{M_{z,2}}{(T-s)^{1/2}} \right) \wedge (M_{z,3}(N+1)).$$

Thanks to BMO tools we have a stability result for quadratic BSDEs (see [20] and [56]):

Proposition 5.9 *Let us assume that (HX0) and (HY0) hold. There exists a constant C that does not depend on N and ε such that*

$$\mathbb{E} \left[\sup_{t \in [0, T]} |Y_t^{N,\varepsilon} - Y_t|^2 \right] + \mathbb{E} \left[\int_0^T |Z_t^{N,\varepsilon} - Z_t|^2 dt \right] \leq C(e_1(N) + e_2(N, \varepsilon))$$

with

$$e_1(N) := \mathbb{E} \left[|g_N(X_T) - g(X_T)|^{2q} \right]^{1/q}, \quad e_2(N, \varepsilon) := \mathbb{E} \left[\left(\int_{T-\varepsilon}^T |f(t, X_t, Y_t^{N,\varepsilon}, Z_t^{N,\varepsilon}) - f(t, X_t, Y_t^{N,\varepsilon}, 0)| dt \right)^{2q} \right]^{1/q},$$

where $q = \frac{r}{r-1}$ and r is defined in Theorem 5.2.

Then, in a second time, we will approximate our modified backward-forward system by a discrete-time one. We will slightly modify the classical discretization by using a non equidistant net with $2n+1$ discretization times. We define the $n+1$ first discretization times on $[0, T-\varepsilon]$ by

$$t_k = T \left(1 - \left(\frac{\varepsilon}{T} \right)^{k/n} \right),$$

and we use an equidistant net on $[T-\varepsilon, T]$ for the last n discretization times:

$$t_k = T - \left(\frac{2n-k}{n} \right) \varepsilon, \quad n \leq k \leq 2n.$$

We denote the time step by $(h_k := t_{k+1} - t_k)_{0 \leq k \leq 2n-1}$. We consider $(X_{t_k}^n)_{0 \leq k \leq 2n}$ the classical Euler scheme for X given by

$$\begin{aligned} X_0^n &= x \\ X_{t_{k+1}}^n &= X_{t_k}^n + h_k b(t_k, X_{t_k}^n) + \sigma(t_k)(W_{t_{k+1}} - W_{t_k}), \quad 0 \leq k \leq 2n-1. \end{aligned} \quad (5.13)$$

²The authors of [56] obtain this result with q^2 instead of q . Nevertheless, we are able to obtain the good result by applying the estimates of [20].

We denote $\rho_s : \mathbb{R}^{1 \times d} \rightarrow \mathbb{R}^{1 \times d}$ the projection on the ball

$$B\left(0, M_{z,1} + \frac{M_{z,2}}{(T-s)^{1/2}}\right)$$

with $M_{z,1}$ and $M_{z,2}$ given by Proposition 5.8. Finally we denote $(Y^{N,\varepsilon,n}, Z^{N,\varepsilon,n})$ our time approximation of $(Y^{N,\varepsilon}, Z^{N,\varepsilon})$. This couple is obtained by a slight modification of the classical dynamic programming equation:

$$\begin{aligned} Y_{t_{2n}}^{N,\varepsilon,n} &= g_N(X_{t_{2n}}^n) \\ Z_{t_k}^{N,\varepsilon,n} &= \rho_{t_{k+1}}\left(\frac{1}{h_k} \mathbb{E}_{t_k}[Y_{t_{k+1}}^{N,\varepsilon,n}(W_{t_{k+1}} - W_{t_k})]\right), \quad 0 \leq k \leq 2n-1, \end{aligned} \quad (5.14)$$

$$Y_{t_k}^{N,\varepsilon,n} = \mathbb{E}_{t_k}[Y_{t_{k+1}}^{N,\varepsilon,n}] + h_k \mathbb{E}_{t_k}[f^\varepsilon(t_k, X_{t_k}^n, Y_{t_{k+1}}^{N,\varepsilon,n}, Z_{t_k}^{N,\varepsilon,n})], \quad 0 \leq k \leq 2n-1, \quad (5.15)$$

where \mathbb{E}_{t_k} stands for the conditional expectation given \mathcal{F}_{t_k} . Let us notice that the classical dynamic programming equation do not use a projection in (5.14): it is the only difference with our time approximation, see e.g. [48] for the classical case. This projection comes directly from the estimate of Z in Proposition 5.8. The aim of our work is to study the error of discretization

$$e(N, \varepsilon, n) := \sup_{0 \leq k \leq 2n} \mathbb{E} \left[\left| Y_{t_k}^{N,\varepsilon,n} - Y_{t_k}^n \right|^2 \right] + \sum_{k=0}^{2n-1} \mathbb{E} \left[\int_{t_k}^{t_{k+1}} \left| Z_{t_k}^{N,\varepsilon,n} - Z_t^n \right|^2 dt \right].$$

It is easy to see that

$$e(N, \varepsilon, n) \leq C(e_1(N) + e_2(N, \varepsilon) + e_3(N, \varepsilon, n)),$$

with $e_1(N)$ and $e_2(N, \varepsilon)$ defined in Proposition 5.9, and

$$e_3(N, \varepsilon, n) := \sup_{0 \leq k \leq 2n} \mathbb{E} \left[\left| Y_{t_k}^{N,\varepsilon,n} - Y_{t_k}^{N,\varepsilon} \right|^2 \right] + \sum_{k=0}^{2n-1} \mathbb{E} \left[\int_{t_k}^{t_{k+1}} \left| Z_{t_k}^{N,\varepsilon,n} - Z_{t_k}^{N,\varepsilon} \right|^2 dt \right].$$

5.4.2 Study of the time approximation error $e_3(N, \varepsilon, n)$

We need an extra assumption.

(HY1). There exists a positive constant $K_{f,t}$ such that $\forall t, t' \in [0, T], \forall x \in \mathbb{R}^d, \forall y \in \mathbb{R}, \forall z \in \mathbb{R}^{1 \times d}$,

$$|f(t, x, y, z) - f(t', x, y, z)| \leq K_{f,t} |t - t'|^{1/2}.$$

Moreover, we set $\varepsilon = Tn^{-a}$ and $N = n^b$, with $a, b \in \mathbb{R}^{+,*}$ two parameters. Before giving our error estimates, we recall two technical lemmas that we will prove in the appendix.

Lemma 5.10 For all constant $M > 0$ there exists a constant C that depends only on T, M and a , such that

$$\prod_{i=0}^{2n-1} (1 + Mh_i) \leq C, \quad \forall n \in \mathbb{N}^*.$$

Lemma 5.11 For all constants $M_1 > 0$ and $M_2 > 0$ there exists a constant C that depends only on T, M_1, M_2 and a , such that

$$\prod_{i=0}^{n-1} \left(1 + M_1 h_i + M_2 \frac{h_i}{T - t_{i+1}} \right) \leq C n^{aM_2}.$$

Firstly, we give a convergence result for the Euler scheme.

Proposition 5.12 Assume (HX0) holds. Then there exists a constant C that does not depend on n , such that

$$\sup_{0 \leq k \leq 2n} \mathbb{E} \left[\left| X_{t_k} - X_{t_k}^n \right|^2 \right] \leq C \frac{\ln n}{n}.$$

Proof. We just have to copy the classical proof to obtain, thanks to Lemma 5.10,

$$\sup_{0 \leq k \leq 2n} \mathbb{E} \left[|X_{t_k} - X_{t_k}^n|^2 \right] \leq C \sup_{0 \leq i \leq 2n-1} h_i = Ch_0.$$

But

$$h_0 = T(1 - n^{-a/n}) \leq C \frac{\ln n}{n},$$

because $(1 - n^{-a/n}) \sim aT \frac{\ln n}{n}$ when $n \rightarrow +\infty$, so the proof is ended. \square

Now, let us treat the BSDE approximation. In a first time we will study the time approximation error on $[T - \varepsilon, T]$.

Proposition 5.13 Assume that (HX0), (HY0) and (HY1) hold. Then there exists a constant C that does not depend on n and such that

$$\sup_{n \leq k \leq 2n} \mathbb{E} \left[|Y_{t_k}^{N,\varepsilon,n} - Y_{t_k}^{N,\varepsilon}|^2 \right] + \sum_{k=n}^{2n-1} \mathbb{E} \left[\int_{t_k}^{t_{k+1}} |Z_{t_k}^{N,\varepsilon,n} - Z_t^{N,\varepsilon}|^2 dt \right] \leq \frac{C \ln n}{n^{1-2b}}.$$

Proof. The BSDE (5.12) has a linear growth with respect to z on $[T - \varepsilon, T]$ so we are allowed to apply classical results which give us that

$$\begin{aligned} \sup_{n \leq k \leq 2n} \mathbb{E} \left[|Y_{t_k}^{N,\varepsilon,n} - Y_{t_k}^{N,\varepsilon}|^2 \right] + \sum_{k=n}^{2n-1} \mathbb{E} \left[\int_{t_k}^{t_{k+1}} |Z_{t_k}^{N,\varepsilon,n} - Z_t^{N,\varepsilon}|^2 dt \right] &\leq C \left(\mathbb{E} \left[|g_N(X_T) - g_N(X_T^n)|^2 \right] + \frac{\varepsilon}{n} \right) \\ &\leq C \left(\frac{\ln n}{n^{1-2b}} + \frac{\varepsilon}{n} \right), \end{aligned}$$

by using the fact that g_N is N -Lipschitz and by applying Proposition 5.12. \square

Remark 5.14

- When $a \geq 1 - 2b$, then $\varepsilon = Tn^{-a} = o(n^{2b-1} \ln n)$. So we do not need to have a discretization grid on $[T - \varepsilon, T]$: $n + 2$ points of discretization are sufficient on $[0, T]$.
- When $a < 1 - 2b$, then is possible to take only $\lceil n^c \rceil$ discretization points on $[T - \varepsilon, T]$ with $a + c = 1 - 2b$. In this case the error bound becomes

$$\sup_{n \leq k \leq 2n} \mathbb{E} \left[|Y_{t_k}^{N,\varepsilon,n} - Y_{t_k}^{N,\varepsilon}|^2 \right] + \sum_{k=n}^{2n-1} \mathbb{E} \left[\int_{t_k}^{t_{k+1}} |Z_{t_k}^{N,\varepsilon,n} - Z_t^{N,\varepsilon}|^2 dt \right] \leq C \left(\frac{\ln n}{n^{1-2b}} + \frac{1}{n^{a+c}} \right),$$

and the Proposition 5.13 stay true.

Now, let us see what happens on $[0, T - \varepsilon]$.

Theorem 5.15 Assume that (HX0), (HY0), (HY1) and (HX1) or (HX1') hold. Then for all $\eta > 0$, there exists a constant C that does not depend on N, ε and n , such that

$$\sup_{0 \leq k \leq 2n} \mathbb{E} \left[|Y_{t_k}^{N,\varepsilon,n} - Y_{t_k}^{N,\varepsilon}|^2 \right] + \sum_{k=0}^{2n-1} \mathbb{E} \left[\int_{t_k}^{t_{k+1}} |Z_{t_k}^{N,\varepsilon,n} - Z_t^{N,\varepsilon}|^2 dt \right] \leq \frac{C}{n^{1-2b-Ka}},$$

with $K = 4(1 + \eta)L_{f,z}^2 M_{z,2}^2$.

Proof. Firstly, we will study the error on Y . From (5.12) and (5.15) we get

$$Y_{t_k}^{N,\varepsilon} - Y_{t_k}^{N,\varepsilon,n} = \mathbb{E}_{t_k} \left[Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n} \right] + \mathbb{E}_{t_k} \int_{t_k}^{t_{k+1}} \left(f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon}) - f(t_k, X_{t_k}^n, Y_{t_{k+1}}^{N,\varepsilon,n}, Z_{t_k}^{N,\varepsilon,n}) \right) ds.$$

We introduce a parameter $\gamma_k > 0$ that will be chosen later. Thanks to Proposition 5.8 and assumption (HY0), f is Lipschitz on $[t_k, t_{k+1}]$ with a Lipschitz constant $K_k := K^1 + \frac{K^2}{(T-t_{k+1})^{1/2}}$ where $K^2 = 2L_{f,z}M_{z,2}$. A combination of Young's inequality $(a+b)^2 \leq (1+\gamma_k h_k)a^2 + (1+\frac{1}{\gamma_k h_k})b^2$ and properties of f gives

$$\begin{aligned} & \mathbb{E} \left| Y_{t_k}^{N,\varepsilon} - Y_{t_k}^{N,\varepsilon,n} \right|^2 \\ & \leq (1+\gamma_k h_k) \mathbb{E} \left| \mathbb{E}_{t_k} \left[Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n} \right] \right|^2 + (1+\eta)^{1/3} K_k^2 \left(h_k + \frac{1}{\gamma_k} \right) \mathbb{E} \int_{t_k}^{t_{k+1}} \left| Z_s^{N,\varepsilon} - Z_{t_k}^{N,\varepsilon,n} \right|^2 ds \\ & \quad + C \left(h_k + \frac{1}{\gamma_k} \right) \left(h_k^2 + \int_{t_k}^{t_{k+1}} \mathbb{E} |X_s - X_{t_k}^n|^2 ds + \int_{t_k}^{t_{k+1}} \mathbb{E} \left| Y_s^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n} \right|^2 ds \right). \end{aligned} \quad (5.16)$$

We define

$$\tilde{Z}_{t_k}^{N,\varepsilon,n} := \frac{1}{h_k} \mathbb{E}_{t_k} [Y_{t_{k+1}}^{N,\varepsilon,n} (W_{t_{k+1}} - W_{t_k})].$$

So, $Z_{t_k}^{N,\varepsilon,n} = \rho_{t_{k+1}}(\tilde{Z}_{t_k}^{N,\varepsilon,n})$. Moreover, Proposition 5.8 implies that $Z_s^{N,\varepsilon} = \rho_{t_{k+1}}(Z_s^{N,\varepsilon})$, and, since $\rho_{t_{k+1}}$ is 1-Lipschitz, we have

$$\left| Z_s^{N,\varepsilon} - Z_{t_k}^{N,\varepsilon,n} \right|^2 = \left| \rho_{t_{k+1}}(Z_s^{N,\varepsilon}) - \rho_{t_{k+1}}(\tilde{Z}_{t_k}^{N,\varepsilon,n}) \right|^2 \leq \left| Z_s^{N,\varepsilon} - \tilde{Z}_{t_k}^{N,\varepsilon,n} \right|^2. \quad (5.17)$$

As in Theorem 5.7, we define $\bar{Z}_{t_k}^{N,\varepsilon}$ by

$$h_k \bar{Z}_{t_k}^{N,\varepsilon} := \mathbb{E}_{t_k} \int_{t_k}^{t_{k+1}} Z_s^{N,\varepsilon} ds = \mathbb{E}_{t_k} \left((Y_{t_{k+1}}^{N,\varepsilon} + \int_{t_k}^{t_{k+1}} f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon}) ds)^t (W_{t_{k+1}} - W_{t_k}) \right).$$

Clearly,

$$\mathbb{E} \int_{t_k}^{t_{k+1}} \left| Z_s^{N,\varepsilon} - \tilde{Z}_{t_k}^{N,\varepsilon,n} \right|^2 ds = \mathbb{E} \int_{t_k}^{t_{k+1}} \left| Z_s^{N,\varepsilon} - \bar{Z}_{t_k}^{N,\varepsilon} \right|^2 ds + h_k \mathbb{E} \left| \bar{Z}_{t_k}^{N,\varepsilon} - \tilde{Z}_{t_k}^{N,\varepsilon,n} \right|^2. \quad (5.18)$$

The Cauchy-Schwartz inequality yields

$$\left| \mathbb{E}_{t_k} \left((Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n})^t (W_{t_{k+1}} - W_{t_k}) \right) \right|^2 \leq h_k \{ \mathbb{E}_{t_k} (|Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}|^2) - \left| \mathbb{E}_{t_k} (Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}) \right|^2 \},$$

and consequently

$$\begin{aligned} h_k \mathbb{E} \left| \bar{Z}_{t_k}^{N,\varepsilon} - \tilde{Z}_{t_k}^{N,\varepsilon,n} \right|^2 & \leq (1+\eta)^{1/3} \mathbb{E} \left[\mathbb{E}_{t_k} (|Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}|^2) - \left| \mathbb{E}_{t_k} (Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}) \right|^2 \right] \\ & \quad + C h_k \mathbb{E} \int_{t_k}^{t_{k+1}} |f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon})|^2 ds. \end{aligned} \quad (5.19)$$

Plugging (5.18) and (5.19) into (5.16), we get:

$$\begin{aligned} \mathbb{E} \left| Y_{t_k}^{N,\varepsilon} - Y_{t_k}^{N,\varepsilon,n} \right|^2 & \leq (1+\gamma_k h_k) \mathbb{E} \left| \mathbb{E}_{t_k} \left[Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n} \right] \right|^2 \\ & \quad + (1+\eta) K_k^2 \left(h_k + \frac{1}{\gamma_k} \right) \mathbb{E} \int_{t_k}^{t_{k+1}} \left| Z_s^{N,\varepsilon} - \bar{Z}_{t_k}^{N,\varepsilon} \right|^2 ds \\ & \quad + C \left(h_k + \frac{1}{\gamma_k} \right) \left(h_k^2 + \int_{t_k}^{t_{k+1}} \mathbb{E} |X_s - X_{t_k}^n|^2 ds + \int_{t_k}^{t_{k+1}} \mathbb{E} \left| Y_s^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n} \right|^2 ds \right) \\ & \quad + (1+\eta)^{2/3} K_k^2 \left(h_k + \frac{1}{\gamma_k} \right) \mathbb{E} \left[\mathbb{E}_{t_k} (|Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}|^2) - \left| \mathbb{E}_{t_k} (Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}) \right|^2 \right] \\ & \quad + C K_k^2 \left(h_k + \frac{1}{\gamma_k} \right) h_k \mathbb{E} \int_{t_k}^{t_{k+1}} |f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon})|^2 ds. \end{aligned}$$

Now write

$$\mathbb{E} \left| Y_s^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n} \right|^2 \leq 2\mathbb{E} \left| Y_s^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon} \right|^2 + 2\mathbb{E} \left| Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n} \right|^2, \quad (5.20)$$

$$\mathbb{E} \left| X_s - X_{t_k}^n \right|^2 \leq 2\mathbb{E} \left| X_s - X_{t_k} \right|^2 + 2\mathbb{E} \left| X_{t_k} - X_{t_k}^n \right|^2, \quad (5.21)$$

we obtain

$$\begin{aligned} \mathbb{E} \left| Y_{t_k}^{N,\varepsilon} - Y_{t_k}^{N,\varepsilon,n} \right|^2 &\leq (1 + \gamma_k h_k) \mathbb{E} \left| \mathbb{E}_{t_k} \left[Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n} \right] \right|^2 \\ &\quad + (1 + \eta) K_k^2 \left(h_k + \frac{1}{\gamma_k} \right) \mathbb{E} \int_{t_k}^{t_{k+1}} \left| Z_s^{N,\varepsilon} - \bar{Z}_{t_k}^{N,\varepsilon} \right|^2 ds \\ &\quad + C \left(h_k + \frac{1}{\gamma_k} \right) \left(h_k^2 + \int_{t_k}^{t_{k+1}} \mathbb{E} \left| X_s - X_{t_k} \right|^2 ds + h_k \mathbb{E} \left| X_{t_k} - X_{t_k}^n \right|^2 \right) \\ &\quad + C \left(h_k + \frac{1}{\gamma_k} \right) \left(\int_{t_k}^{t_{k+1}} \mathbb{E} \left| Y_s^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon} \right|^2 ds + h_k \mathbb{E} \left| Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n} \right|^2 \right) \\ &\quad + (1 + \eta)^{2/3} K_k^2 \left(h_k + \frac{1}{\gamma_k} \right) \mathbb{E} \left[\mathbb{E}_{t_k} \left(\left| Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n} \right|^2 \right) - \left| \mathbb{E}_{t_k} (Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}) \right|^2 \right] \\ &\quad + C K_k^2 \left(h_k + \frac{1}{\gamma_k} \right) h_k \mathbb{E} \int_{t_k}^{t_{k+1}} \left| f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon}) \right|^2 ds. \end{aligned}$$

Taking $\gamma_k = (1 + \eta)^{2/3} K_k^2$: for h_k small enough, it gives

$$\begin{aligned} \mathbb{E} \left| Y_{t_k}^{N,\varepsilon} - Y_{t_k}^{N,\varepsilon,n} \right|^2 &\leq (1 + Ch_k + (1 + \eta)^{2/3} K_k^2 h_k) \mathbb{E} \left| Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n} \right|^2 + Ch_k^2 + Ch_k \max_{0 \leq k \leq n} \mathbb{E} \left| X_{t_k} - X_{t_k}^n \right|^2 \\ &\quad + C \mathbb{E} \int_{t_k}^{t_{k+1}} \left| Z_s^{N,\varepsilon} - \bar{Z}_{t_k}^{N,\varepsilon} \right|^2 ds + C \int_{t_k}^{t_{k+1}} \mathbb{E} \left| X_s - X_{t_k} \right|^2 ds \\ &\quad + C \int_{t_k}^{t_{k+1}} \mathbb{E} \left| Y_s^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon} \right|^2 ds + Ch_k \mathbb{E} \int_{t_k}^{t_{k+1}} \left| f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon}) \right|^2 ds, \end{aligned}$$

because $K_k^2 h_k \leq C(h_0 + h_k(T - t_{k+1})^{-1}) \leq C \frac{\ln n}{n}$. The Gronwall's lemma gives us

$$\begin{aligned} \mathbb{E} \left| Y_{t_k}^{N,\varepsilon} - Y_{t_k}^{N,\varepsilon,n} \right|^2 &\leq C \sum_{j=0}^{n-1} \left[\prod_{i=0}^{j-1} (1 + Ch_i + (1 + \eta)^{2/3} K_i^2 h_i) \right] \left[h_j^2 + h_j \max_{0 \leq l \leq n} \mathbb{E} \left| X_{t_l} - X_{t_l}^n \right|^2 \right. \\ &\quad + \mathbb{E} \int_{t_j}^{t_{j+1}} \left(\left| Z_s^{N,\varepsilon} - \bar{Z}_{t_j}^{N,\varepsilon} \right|^2 + \left| X_s - X_{t_j} \right|^2 + \left| Y_s^{N,\varepsilon} - Y_{t_{j+1}}^{N,\varepsilon} \right|^2 \right) ds \\ &\quad + h_j \mathbb{E} \int_{t_j}^{t_{j+1}} \left| f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon}) \right|^2 ds \left. \right] \\ &\quad + \left[\prod_{i=0}^{n-1} (1 + Ch_i + (1 + \eta)^{2/3} K_i^2 h_i) \right] \mathbb{E} \left| Y_{t_n}^{N,\varepsilon} - Y_{t_n}^{N,\varepsilon,n} \right|^2. \end{aligned}$$

Then, we apply Lemma 5.11:

$$\begin{aligned} \mathbb{E} \left| Y_{t_k}^{N,\varepsilon} - Y_{t_k}^{N,\varepsilon,n} \right|^2 &\leq C n^{(1+\eta)(K^2)^2 a} \left[h_0 + \max_{0 \leq l \leq n} \mathbb{E} \left| X_{t_l} - X_{t_l}^n \right|^2 \right. \\ &\quad + \sum_{j=0}^n \mathbb{E} \left(\int_{t_j}^{t_{j+1}} \left| Z_s^{N,\varepsilon} - \bar{Z}_{t_j}^{N,\varepsilon} \right|^2 + \left| X_s - X_{t_j} \right|^2 + \left| Y_s^{N,\varepsilon} - Y_{t_{j+1}}^{N,\varepsilon} \right|^2 ds \right) \\ &\quad + h_0 \mathbb{E} \int_0^{t_n} \left| f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon}) \right|^2 ds + \mathbb{E} \left| Y_{t_n}^{N,\varepsilon} - Y_{t_n}^{N,\varepsilon,n} \right|^2 \left. \right]. \end{aligned}$$

A classical estimation gives us $\mathbb{E} \left[|X_s - X_{t_j}|^2 \right] \leq C |s - t_j|$. Moreover, since $Z^{N,\varepsilon}$ is bounded,

$$\begin{aligned} \mathbb{E} \int_0^{t_n} |f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon})|^2 ds &\leq CT(1 + |Y^{N,\varepsilon}|_\infty) + C\mathbb{E} \left[\int_0^{t_n} |Z_s^{N,\varepsilon}|^4 ds \right] \\ &\leq CT(1 + |Y^{N,\varepsilon}|_\infty) + Cn^{2b} \mathbb{E} \left[\int_0^T |Z_s^{N,\varepsilon}|^2 ds \right]. \end{aligned}$$

But we have an a priori estimate for $\mathbb{E} \left[\int_0^T |Z_s^{N,\varepsilon}|^2 ds \right]$ that does not depend on N and ε . So

$$\mathbb{E} \int_0^{t_n} |f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon})|^2 ds \leq Cn^{2b}. \quad (5.22)$$

With the same type of argument we also have

$$\mathbb{E} \left| Y_s^{N,\varepsilon} - Y_{t_{j+1}}^{N,\varepsilon} \right|^2 \leq Ch_j n^{2b}. \quad (5.23)$$

If we add Zhang's path regularity theorem 5.7, Proposition 5.12 and Proposition 5.13, we finally obtain

$$\mathbb{E} \left| Y_{t_k}^{N,\varepsilon} - Y_{t_k}^{N,\varepsilon,n} \right|^2 \leq Cn^{(1+\eta)(K^2)^2a} \frac{n^{2b} \ln n}{n} = C \frac{\ln n}{n^{1-2b-(1+\eta)(K^2)^2a}}. \quad (5.24)$$

Now, let us deal with the error on Z . First of all, (5.17) gives us

$$\sum_{k=0}^{n-1} \mathbb{E} \left[\int_{t_k}^{t_{k+1}} |Z_{t_k}^{N,\varepsilon,n} - Z_t^{N,\varepsilon}|^2 dt \right] \leq \sum_{k=0}^{n-1} \mathbb{E} \left[\int_{t_k}^{t_{k+1}} |\tilde{Z}_{t_k}^{N,\varepsilon,n} - Z_t^{N,\varepsilon}|^2 dt \right].$$

For $0 \leq k \leq n-1$, we can use (5.18) and (5.19) to obtain

$$\begin{aligned} \mathbb{E} \left[\int_{t_k}^{t_{k+1}} |\tilde{Z}_{t_k}^{N,\varepsilon,n} - Z_t^{N,\varepsilon}|^2 dt \right] &\leq \mathbb{E} \left[\int_{t_k}^{t_{k+1}} |\bar{Z}_{t_k}^{N,\varepsilon} - Z_t^{N,\varepsilon}|^2 dt \right] \\ &\quad + (1+\eta)^{2/3} \mathbb{E} \left[\mathbb{E}_{t_k} (|Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}|^2) - \left| \mathbb{E}_{t_k} (Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}) \right|^2 \right] \\ &\quad + Ch_k \mathbb{E} \left[\int_{t_k}^{t_{k+1}} |f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon})|^2 ds \right]. \end{aligned}$$

Inequality (5.22) and estimates for Z give us

$$\begin{aligned} &\sum_{k=0}^{n-1} \mathbb{E} \left[\int_{t_k}^{t_{k+1}} |Z_{t_k}^{N,\varepsilon,n} - Z_t^{N,\varepsilon}|^2 dt \right] \\ &\leq \sum_{k=0}^{n-1} \mathbb{E} \left[\int_{t_k}^{t_{k+1}} |\bar{Z}_{t_k}^{N,\varepsilon} - Z_t^{N,\varepsilon}|^2 dt \right] \end{aligned} \quad (5.25)$$

$$\begin{aligned} &\quad + (1+\eta)^{2/3} \sum_{k=0}^{n-1} \mathbb{E} \left[\mathbb{E}_{t_k} (|Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}|^2) - \left| \mathbb{E}_{t_k} (Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}) \right|^2 \right] \\ &\quad + Ch_0 \mathbb{E} \left[\int_0^T |f(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon})|^2 ds \right] \\ &\leq \sum_{k=0}^{n-1} \mathbb{E} \left[\int_{t_k}^{t_{k+1}} |\bar{Z}_{t_k}^{N,\varepsilon} - Z_t^{N,\varepsilon}|^2 dt \right] \\ &\quad + (1+\eta)^{2/3} \sum_{k=0}^{n-1} \mathbb{E} \left[\mathbb{E}_{t_k} (|Y_{t_k}^{N,\varepsilon} - Y_{t_k}^{N,\varepsilon,n}|^2) - \left| \mathbb{E}_{t_k} (Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}) \right|^2 \right] \\ &\quad + C\mathbb{E} \left[|Y_{t_n}^{N,\varepsilon} - Y_{t_n}^{N,\varepsilon,n}|^2 \right] + Ch_0 n^{2b}, \end{aligned} \quad (5.26)$$

with an index change in the penultimate line. Then, by using (5.16) we get

$$\begin{aligned}
& (1+\eta)^{2/3} \mathbb{E} \left[\mathbb{E}_{t_k} \left(\left| Y_{t_k}^{N,\varepsilon} - Y_{t_k}^{N,\varepsilon,n} \right|^2 \right) - \left| \mathbb{E}_{t_k} (Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}) \right|^2 \right] \\
& \leq C \gamma_k h_k \mathbb{E} \left[\mathbb{E}_{t_k} \left[Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n} \right]^2 \right] + (1+\eta) K_k^2 \left(h_k + \frac{1}{\gamma_k} \right) \mathbb{E} \int_{t_k}^{t_{k+1}} \left| Z_s^{N,\varepsilon} - Z_{t_k}^{N,\varepsilon,n} \right|^2 ds \\
& \quad + C \left(h_k + \frac{1}{\gamma_k} \right) h_k \left(h_k + \sup_{s \in [t_k, t_{k+1}]} \mathbb{E} \left[\left| X_s - X_{t_k}^n \right|^2 + \left| Y_s^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n} \right|^2 \right] \right). \tag{5.27}
\end{aligned}$$

Thanks to (5.20), (5.21), (5.23) and a classical estimation on $\mathbb{E} \left[\left| X_s - X_{t_k}^n \right|^2 \right]$ we have

$$\sup_{s \in [t_k, t_{k+1}]} \mathbb{E} \left[\left| X_s - X_{t_k}^n \right|^2 + \left| Y_s^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n} \right|^2 \right] \leq C \left(h_k n^{2b} + \mathbb{E} \left[\left| Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n} \right|^2 \right] \right).$$

Let us set $\gamma_k = 3(1+\eta)K_k^2$. We recall that $h_k K_k^2 \leq \frac{C \ln n}{n} \rightarrow 0$ when $n \rightarrow 0$. So, for n big enough, (5.27) becomes

$$\begin{aligned}
& (1+\eta)^{2/3} \mathbb{E} \left[\mathbb{E}_{t_k} \left(\left| Y_{t_k}^{N,\varepsilon} - Y_{t_k}^{N,\varepsilon,n} \right|^2 \right) - \left| \mathbb{E}_{t_k} (Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n}) \right|^2 \right] \\
& \leq \frac{C \ln n}{n} \mathbb{E} \left[\left| Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n} \right|^2 \right] + \frac{1}{2} \mathbb{E} \int_{t_k}^{t_{k+1}} \left| Z_s^{N,\varepsilon} - Z_{t_k}^{N,\varepsilon,n} \right|^2 ds \\
& \quad + C h_0 h_k n^{2b}.
\end{aligned}$$

If we inject this last estimate in (5.26) and we use Theorem 5.7, we obtain

$$\frac{1}{2} \sum_{k=0}^{n-1} \mathbb{E} \left[\int_{t_k}^{t_{k+1}} \left| Z_t^{N,\varepsilon,n} - Z_t^{N,\varepsilon} \right|^2 dt \right] \leq C h_0 n^{2b} + C \ln n \sup_{0 \leq k \leq n-1} \mathbb{E} \left[\left| Y_{t_{k+1}}^{N,\varepsilon} - Y_{t_{k+1}}^{N,\varepsilon,n} \right|^2 \right].$$

By using (5.24) and Proposition 5.13, we finally have

$$\sup_{0 \leq k \leq 2n} \mathbb{E} \left[\left| Y_{t_k}^{N,\varepsilon,n} - Y_{t_k}^{N,\varepsilon} \right|^2 \right] + \sum_{k=0}^{2n-1} \mathbb{E} \left[\int_{t_k}^{t_{k+1}} \left| Z_t^{N,\varepsilon,n} - Z_t^{N,\varepsilon} \right|^2 dt \right] \leq C \frac{(\ln n)^2}{n^{1-2b-Ka}},$$

with $K = 4(1+\eta)L_{f,z}^2 M_{z,2}^2$. Since this estimate is true for every $\eta > 0$, we have proved the result. \square

5.4.3 Study of the global error $e(N, \varepsilon, n)$

Let us study errors $e_1(N)$ and $e_2(N, \varepsilon)$.

Proposition 5.16 *Let us assume that (HX0) and (HY0) hold. There exists a constant $C > 0$ such that*

$$e_2(N, \varepsilon) \leq \frac{C}{n^{2a-4b}}.$$

Proof. We just have to notice that $\left| f(t, X_t, Y_t^{N,\varepsilon}, Z_t^{N,\varepsilon}) - f(t, X_t, Y_t^{N,\varepsilon}, 0) \right| \leq C \left| Z_t^{N,\varepsilon} \right|^2$ and $\left| Z_t^{N,\varepsilon} \right|$ is bounded by Cn^b . \square

For g_N we use the classical Lipschitz approximation

$$g_N(x) = \inf \{ g(u) + N |x - u| \mid u \in \mathbb{R}^d \}.$$

Proposition 5.17 *We assume that (HX0) holds and g is α -Hölder. Then, there exists a constant C such that*

$$e_1(N) \leq \frac{C}{n^{\frac{2b\alpha}{1-\alpha}}}.$$

Proof. g_N is a N -Lipschitz function and $g_N \rightarrow g$ when $N \rightarrow +\infty$ uniformly on \mathbb{R}^d . More precisely, we have

$$|g - g_N|_\infty \leq \frac{C}{N^{\frac{\alpha}{1-\alpha}}}.$$

□

Remark 5.18 For some explicit examples, it is possible to have a better convergence speed. For example, let us take $g(x) = (|x|^\alpha \mathbb{1}_{x \geq 0}) \wedge C$ and assume that σ is invertible. Then, we can use the fact that this function is not Lipschitz only in 0, and obtain

$$e_1(N) \leq \frac{C}{n^{\frac{2\alpha b}{1-\alpha}}} \mathbb{P}\left(X_T \in \left[0, N^{\frac{-1}{1-\alpha}}\right]\right)^{1/q} \leq \frac{C}{n^{\frac{b}{1-\alpha}(2\alpha + \frac{1}{q})}}.$$

Remark 5.19 It is also possible to obtain convergence speed when g is not α -Hölder. For example, we assume that σ is invertible and we set $g(x) = \prod_{i=1}^d \mathbb{1}_{x_i > 0}(x)$. Then

$$e_1(N) \leq C \left[\sum_{i=1}^d \mathbb{P}((X_T)_i \in [0, 1/N]) \right]^{1/q} \leq \frac{C}{N^{1/q}} = \frac{C}{n^{b/q}}.$$

Now we are able to gather all these errors.

Theorem 5.20 We assume that (HX0), (HY0), (HY1), and (HX1) or (HX1') hold. We assume also that g is α -Hölder. Then for all $\eta > 0$, there exists a constant $C > 0$ that does not depend on n such that

$$e(n) := e(N, \varepsilon, n) \leq \frac{C}{n^{\frac{2\alpha}{(2-\alpha)(2+K)-2+2\alpha}}},$$

with $K = 4(1 + \eta)L_{f,z}^2 M_{z,2}^2$.

Proof. Thanks to Theorem 5.15, Proposition 5.16 and Proposition 5.17 we have

$$e(n) \leq \frac{C}{n^{1-2b-Ka}} + \frac{C}{n^{2a-4b}} + \frac{C}{n^{\frac{2\alpha b}{1-\alpha}}}.$$

Then we only need to set $a := \frac{1+2b}{2+K}$ and $b := \frac{1-\alpha}{(2-\alpha)(2+K)-2+2\alpha}$ to obtain the result. □

Theorem 5.21 We assume that assumptions of Theorem 5.20 hold. Moreover we assume that f has a sub-quadratic growth with respect to z : there exists $0 < \beta < 1$ such that, for all $t \in [0, T]$, $x \in \mathbb{R}^d$, $y \in R$, $z, z' \in \mathbb{R}^{1 \times d}$,

$$|f(t, x, y, z) - f(t, x, y, z')| \leq (K_{f,z} + L_{f,z}(|z|^\beta + |z'|^\beta)) |z - z'|.$$

Then, for all $\eta > 0$, there exists a constant $C > 0$ that does not depend on n such that

$$e(n) \leq \frac{C}{n^{\alpha-\eta}}.$$

Proof. Firstly, Proposition 5.8 stays true. Then, let us remark that for all $\delta > 0$ we have

$$|f(t, x, y, z) - f(t, x, y, z')| \leq (C_\delta + \delta L_{f,z}(|z| + |z'|)) |z - z'|.$$

So, Theorem 5.15 also stays true with a constant C that depends on δ and $K_\delta = 4(1 - \eta)\delta^2 L_{f,z}^2 M_{z,2}^2$. Finally, like in Theorem 5.20, we obtain that

$$e(n) \leq C_\delta n^{-\frac{2\alpha}{(2-\alpha)(2+K_\delta)-2+2\alpha}}.$$

To conclude, we just have to remark that

$$\lim_{\delta \rightarrow 0^+} \frac{2\alpha}{(2-\alpha)(2+K_\delta)-2+2\alpha} = \alpha^-.$$

□

Remark 5.22 We are able to specify Remark 5.14 in our case, when $a = \frac{1+2b}{2+K}$ and $b = \frac{1-\alpha}{(2-\alpha)(2+K)-2+2\alpha}$.

- When $K \leq \frac{2-3\alpha}{2-\alpha}$, that is to say, when $\alpha < 2/3$ and K is sufficiently small, then we do not need to have a discretization grid on $[T - \varepsilon, T]$.
- When $K > \frac{2-3\alpha}{2-\alpha}$, then it is possible to take only $\lceil n^c \rceil$ discretization points on $[T - \varepsilon, T]$ with

$$c = 1 + \frac{3\alpha - 4}{(2 - \alpha)(2 + K) - 2 + 2\alpha}.$$

Theorem 5.20 is not interesting in practice because the speed of convergence depends strongly on K . But, we just see that the global error becomes $e(n) \leq \frac{C}{n^{\alpha-\eta}}$ when we are allowed to choose K as small as we want. Under extra assumption we can show that we are allowed to take the constant $M_{z,2}$ as small as we want.

(HX2). b is bounded on $[0, T] \times \mathbb{R}^d$ by a constant M_b .

Theorem 5.23 We assume that (HX0), (HY0), (HY1), (HX2) and (HX1) or (HX1') hold. We assume also that g is α -Hölder. Then for all $\eta > 0$, there exists a constant $C > 0$ that does not depend on n such that

$$e(n) \leq \frac{C}{n^{\alpha-\eta}}.$$

Remark 5.24 With the assumptions of the previous theorem, it is also possible to have an estimate of the global error for examples given in Remarks 5.18 and 5.19. When $g(x) = (|x|^\alpha \mathbb{1}_{x \geq 0}) \wedge C$, we have

$$e(n) \leq \frac{C}{n^{\alpha + \frac{1-\alpha}{1+2q} - \eta}},$$

and when $g(x) = \prod_{i=1}^d \mathbb{1}_{x_i > 0}(x)$, we have

$$e(n) \leq \frac{C}{n^{\frac{1}{1+2q} - \eta}}.$$

Proof. Firstly, we suppose that f is differentiable with respect to z . Thanks to Remark 5.5 we see that it is sufficient to show that

$$\mathbb{E}^{\mathbb{Q}^{N,\varepsilon}} \left[\int_t^T |Z_s^{N,\varepsilon}|^2 ds \middle| \mathcal{F}_t \right]$$

is small uniformly in ω , N and ε when t is close to T . We will obtain an estimation for this quantity by applying the same computation as [20] for the BMO norm estimate of Z page 831. Thus we have

$$\mathbb{E}^{\mathbb{Q}^{N,\varepsilon}} \left[\int_t^T |Z_s^{N,\varepsilon}|^2 ds \middle| \mathcal{F}_t \right] \leq \mathbb{E}^{\mathbb{Q}^{N,\varepsilon}} \left[\left| \varphi(Y_T^{N,\varepsilon}) - \varphi(Y_t^{N,\varepsilon}) \right| \middle| \mathcal{F}_t \right] + C(T - t),$$

with $\varphi(x) = (e^{2c(x+m)} - 2c(x+m) - 1)/(2c^2)$, $m = |Y|_\infty$ and c that depends on constants in assumption (HY0) but does not depend on $\nabla_z f$. Let us notice that m , c and so φ do not depend on N and ε . Since Y is bounded, φ is a Lipschitz function, so

$$\mathbb{E}^{\mathbb{Q}^{N,\varepsilon}} \left[\int_t^T |Z_s^{N,\varepsilon}|^2 ds \middle| \mathcal{F}_t \right] \leq C \mathbb{E}^{\mathbb{Q}^{N,\varepsilon}} \left[\left| Y_T^{N,\varepsilon} - Y_t^{N,\varepsilon} \right| \middle| \mathcal{F}_t \right] + C(T - t).$$

We denote by $(Y^{N,\varepsilon,t,x}, Z^{N,\varepsilon,t,x})$ the solution of BSDE (5.12) when $X_t^{t,x} = x$. As usual, we set $X_s^{t,x} = x$ and $Z_s^{N,\varepsilon,t,x} = 0$ for $s \leq t$ and we define $u^{N,\varepsilon}(t, x) := Y_t^{N,\varepsilon,t,x}$. Then we give a proposition that we will prove in the appendix.

Proposition 5.25 We assume that (HX0), (HY0), (HY1), (HX2) and (HX1) or (HX1') hold. We assume also that g is uniformly continuous on \mathbb{R}^d . Then $u^{N,\varepsilon}$ is uniformly continuous on $[0, T] \times \mathbb{R}^d$ and there exists ω a concave modulus of continuity for all functions in $\{u^{N,\varepsilon} | N \in \mathbb{N}, \varepsilon > 0\}$: i.e. ω does not depend on N and ε .

Then

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}^{N,\varepsilon}} \left[\left| Y_T^{N,\varepsilon} - Y_t^{N,\varepsilon} \right| \middle| \mathcal{F}_t \right] &= \mathbb{E}^{\mathbb{Q}^{N,\varepsilon}} \left[\left| u^{N,\varepsilon}(T, X_T) - u^{N,\varepsilon}(t, X_t) \right| \middle| \mathcal{F}_t \right] \\
&\leq \mathbb{E}^{\mathbb{Q}^{N,\varepsilon}} \left[\mathbb{1}_{\left| \int_t^T \sigma(s) d\tilde{W}_s \right| \leq \nu} \left| u^{N,\varepsilon}(T, X_T) - u^{N,\varepsilon}(t, X_t) \right| \right. \\
&\quad \left. + 2 \left| Y^{N,\varepsilon} \right|_\infty \mathbb{1}_{\left| \int_t^T \sigma(s) d\tilde{W}_s \right| > \nu} \middle| \mathcal{F}_t \right] \\
&\leq \mathbb{E}^{\mathbb{Q}^{N,\varepsilon}} \left[\omega \left(|T-t| + \mathbb{1}_{\left| \int_t^T \sigma(s) d\tilde{W}_s \right| \leq \nu} |X_T - X_t| \right) \right. \\
&\quad \left. + 2 \left| Y^{N,\varepsilon} \right|_\infty \mathbb{1}_{\left| \int_t^T \sigma(s) d\tilde{W}_s \right| > \nu} \middle| \mathcal{F}_t \right],
\end{aligned}$$

with $d\tilde{W}_s = dW_s - \nabla_z f^\varepsilon(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon}) ds$. But,

$$\begin{aligned}
&\mathbb{1}_{\left| \int_t^T \sigma(s) d\tilde{W}_s \right| \leq \nu} |X_T - X_t| \\
&= \mathbb{1}_{\left| \int_t^T \sigma(s) d\tilde{W}_s \right| \leq \nu} \left| \int_t^T b(s, X_s) ds + \int_t^T \nabla_z f^\varepsilon(s, X_s, Y_s^{N,\varepsilon}, Z_s^{N,\varepsilon}) ds + \int_t^T \sigma(s) d\tilde{W}_s \right| \\
&\leq M_b(T-t) + \nu + C \int_t^T (1 + |Z_s^{N,\varepsilon}|) ds \\
&\leq C(T-t) + \nu + C(T-t)^{1/2} \left(\int_t^T |Z_s^{N,\varepsilon}|^2 ds \right)^{1/2}.
\end{aligned}$$

Since ω is concave, we have by Jensen's inequality

$$\begin{aligned}
&\mathbb{E}^{\mathbb{Q}^{N,\varepsilon}} \left[\omega \left(|T-t| + \mathbb{1}_{\left| \int_t^T \sigma(s) d\tilde{W}_s \right| \leq \nu} |X_T - X_t| \right) \middle| \mathcal{F}_t \right] \\
&\leq \omega \left(C|T-t| + \nu + C(T-t)^{1/2} \mathbb{E}^{\mathbb{Q}^{N,\varepsilon}} \left[\left(\int_t^T |Z_s^{N,\varepsilon}|^2 ds \right)^{1/2} \middle| \mathcal{F}_t \right] \right) \\
&\leq \omega \left(C|T-t| + \nu + C(T-t)^{1/2} \mathbb{E}^{\mathbb{Q}^{N,\varepsilon}} \left[\int_t^T |Z_s^{N,\varepsilon}|^2 ds \middle| \mathcal{F}_t \right]^{1/2} \right) \\
&\leq \omega \left(C|T-t| + \nu + C(T-t)^{1/2} \|Z^{N,\varepsilon}\|_{BMO(\mathbb{Q})} \right).
\end{aligned}$$

But, $\|Z^{N,\varepsilon}\|_{BMO(\mathbb{Q})}$ only depends on constants in assumption (HY0), so it is bounded uniformly in N and ε . Moreover, $\left| \int_t^T \sigma(s) d\tilde{W}_s \right|$ is independent of \mathcal{F}_t so we have by the Markov inequality

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}^{N,\varepsilon}} \left[\mathbb{1}_{\left| \int_t^T \sigma(s) d\tilde{W}_s \right| > \nu} \middle| \mathcal{F}_t \right] &= \mathbb{Q}^{N,\varepsilon} \left(\left| \int_t^T \sigma(s) d\tilde{W}_s \right| > \nu \right) \\
&\leq \frac{C(T-t)^{1/2}}{\nu}.
\end{aligned}$$

Finally, we have

$$\begin{aligned}
\mathbb{E}^{\mathbb{Q}^{N,\varepsilon}} \left[\left| Y_T^{N,\varepsilon} - Y_t^{N,\varepsilon} \right| \middle| \mathcal{F}_t \right] &\leq \omega \left(C|T-t|^{1/2} + \nu \right) + C \frac{(T-t)^{1/2}}{\nu} \\
&\leq \omega \left(C|T-t|^{1/2} + |T-t|^{1/4} \right) + C|T-t|^{1/4},
\end{aligned}$$

by setting $\nu = |T-t|^{1/4}$, and $\mathbb{E}^{\mathbb{Q}^{N,\varepsilon}} \left[\left| Y_T^{N,\varepsilon} - Y_t^{N,\varepsilon} \right| \middle| \mathcal{F}_t \right] \rightarrow 0$ uniformly in ω , N and ε when $t \rightarrow T$. When f is not differentiable with respect to z but is only locally Lipschitz then we can prove the result by a standard approximation. \square

5.5 Some additional results on the time dependent estimate of Z^3

5.5.1 What happens if σ does not depend on time ?

We have seen that the key point of our approximation results is the time dependent estimate of Z in Theorem 5.4. This estimate needs the technical assumption (HX1). When σ does not depend on time, this assumption becomes

(HX1). b is differentiable with respect to x and there exists $\lambda \in \mathbb{R}^+$ such that for all $\eta \in \mathbb{R}^d$

$$\left| {}^t\eta \sigma^t \sigma^t \nabla b(s, x) \eta \right| \leq \lambda \left| {}^t\eta \sigma \right|^2. \quad (5.28)$$

It is possible to obtain some equivalent assumptions.

Proposition 5.26 *Let us assume that (HX0) hold, b is differentiable with respect to x and σ does not depend on time. Then, following assertions are equivalent:*

- (i) $\forall x \in \mathbb{R}^d, \forall s \in [0, T], \text{Im } \nabla b(s, x) \sigma \subset \text{Im } \sigma$,
- (ii) $\forall x \in \mathbb{R}^d, \forall s \in [0, T], \ker {}^t\sigma \subset \ker {}^t\sigma^t \nabla b(s, x)$,
- (iii) $\exists f : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ bounded such that $\forall x \in \mathbb{R}^d, \forall s \in [0, T], \nabla b(s, x) \sigma = \sigma f(s, x)$,
- (iv) there exist $A : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ and $B : [0, T] \rightarrow \mathbb{R}^{d \times d}$ such that A is differentiable with respect to x , ∇A is bounded and $\forall x \in \mathbb{R}^d, \forall s \in [0, T], b(s, x) \sigma = \sigma A(s, x) + B(s)$,
- (v) there exists $\lambda \in \mathbb{R}^+$ such that $\forall \eta \in \mathbb{R}^d$

$$\left| {}^t\eta \sigma^t \sigma^t \nabla b(s, x) \eta \right| \leq \lambda \left| {}^t\eta \sigma \right|^2. \quad (5.29)$$

Proof.

(iii) \Rightarrow (i), (iii) \Rightarrow (ii), (iii) \Rightarrow (v) and (iii) \Leftrightarrow (iv): Trivial.

(i) \Rightarrow (iii): There exists $\tilde{\sigma}$ such that $\sigma \tilde{\sigma}|_{\text{Im } \sigma} = Id|_{\text{Im } \sigma}$. Since $\text{Im } \nabla b(s, x) \sigma \subset \text{Im } \sigma$ then $\nabla b(s, x) \sigma = \sigma \tilde{\sigma} \nabla b(s, x) \sigma := \sigma f(s, x)$. ∇b is bounded so f is also bounded.

(ii) \Rightarrow (iii): The proof is the same when we notice that there exists $\tilde{\sigma}$ such that $\tilde{\sigma} \sigma|_M = Id|_M$ with $M \oplus \ker \sigma = \mathbb{R}^d$.

(v) \Rightarrow (ii): Let $S_d^+(\mathbb{R})$ denote the space of non negative symmetric matrix. Since $\sigma^t \sigma \in S_d^+(\mathbb{R})$, there exists $s \in S_d^+(\mathbb{R})$ such that $\sigma^t \sigma = s^2$. It is easy to show that $\ker s = \ker {}^t\sigma$ and $\ker s^t \nabla b(s, x) = \ker {}^t\sigma^t \nabla b(s, x)$. Moreover, $\forall \eta \in \mathbb{R}^d$ and $|{}^t\eta \sigma|^2 = |{}^t\eta s|^2$, so we are allowed to assume that $\sigma \in S_d^+(\mathbb{R})$.

If σ is invertible or $\sigma = 0$, the result is proved. Otherwise, let us consider η_1 and η_2 two eigenvectors of σ linked with eigenvalues $\lambda_1 \neq 0$ and $\lambda_2 = 0$. (v) gives us, for all $\alpha \in \mathbb{R}$,

$$\begin{aligned} \left| {}^t(\eta_1 + \alpha \eta_2) \sigma^{2t} \nabla b(s, x) (\eta_1 + \alpha \eta_2) \right| &= \left| \lambda_1^{2t} \eta_1^t \nabla b(s, x) \eta_1 + \alpha \lambda_1^{2t} \eta_1^t \nabla b(s, x) \eta_2 \right| \\ &\leq \lambda \lambda_1^2 |\eta_1|^2. \end{aligned}$$

$\left| {}^t(\eta_1 + \alpha \eta_2) \sigma^{2t} \nabla b(s, x) (\eta_1 + \alpha \eta_2) \right|$ is bounded with respect to α , implying ${}^t\eta_1^t \nabla b(s, x) \eta_2 = 0$.

Since $\sigma \in S_d^+(\mathbb{R})$, we have $\ker \sigma \oplus \text{Im } \sigma$. So, for all $\eta_2 \in \ker \sigma$, we have shown ${}^t\eta_2^t \nabla b(s, x) \eta_2 \in (\text{Im } \sigma)^\perp = \ker \sigma$, that is to say, $\sigma^t \nabla b(s, x) \eta_2 = 0$. This last equality allows us to conclude that $\ker \sigma \subset \ker \sigma^t \nabla b(s, x)$. \square

The assertion (v) allows us to reduce assumptions on the regularity of b . Indeed, when A is only a Lipschitz function with respect to x with a Lipschitz constant K_A , then it is possible to do a classical approximation of A by a sequence of differentiable functions $(A_n)_{n \in \mathbb{N}}$ such that $|\nabla A_n(s, x)| \leq K_A$.

³This section has been added after the submission of the article.

(HX1''). σ does not depend on time. There exist $A : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^{d \times d}$ and $B : [0, T] \rightarrow \mathbb{R}^{d \times d}$ such that $\forall x \in \mathbb{R}^d, \forall s \in [0, T], b(s, x)\sigma = \sigma A(s, x) + B(s)$.

Then we obtain a new version of Theorem 5.4.

Theorem 5.27 *Suppose that (HX0), (HY0) and (HX1'') hold. Moreover, suppose that g is lower (or upper) semi-continuous. Then there exists a version of Z and there exist two constants $C, C' \in \mathbb{R}^+$ that depend only in $T, M_g, M_f, K_{f,x}, K_{f,y}, K_{f,z}$ and $L_{f,z}$ such that, $\forall t \in [0, T[$,*

$$|Z_t| \leq C + C'(T - t)^{-1/2}.$$

5.5.2 Some examples and counterexamples

As we already explain in the introduction, Theorem 5.4 is also interesting for BSDEs with a linear growth with respect to z . Indeed, such an estimation has been already proved when σ is invertible by using the Bismut-Elworthy formula, see e.g. [44], while in our case we do not need to assume the invertibility of σ . Nevertheless, technical assumptions (HX1) or (HX1'') appear. Let us see what could happen when these assumptions do not hold. The most simple counterexample is given by Zhang in [85]. Let us consider our SDE in dimension 1. We take:

- $T = 2$,
- $b = 0$,
- $\forall t \in [0, 2], \sigma(t) = (1 - t)\mathbb{1}_{t < 1}$,
- $f = 0$,
- $\forall x \in \mathbb{R}, g(x) = \arctan\left(\frac{x}{|x|^{3/4}}\right)$.

Proposition 5.28 $\forall M > 0, \exists 0 \leq t_1 < t_2 < 1, \mathbb{P}(|Z_{[t_1, t_2]}| \geq M) > 0$.

Proof. Let us denote $u(t, x) := Y_t^{t,x}$ with $(Y_t^{t,x}, Z_t^{t,x})$ the solution of the BSDE (5.5) where X verifies $X_t = x$. For $t \in [1, 2]$, obviously $u(t, x) = g(x)$ and $Z_t^{t,x} = 0$. For $t \in [0, 1[$, we have

$$\begin{aligned} u(t, x) &= \mathbb{E} \left[g \left(x + \int_t^1 (1 - s) dW_s \right) \right] \\ &= \frac{1}{\sqrt{2\pi\sigma_t}} \int_{\mathbb{R}} g(y) \exp \left(-\frac{(y - x)^2}{2\sigma_t^2} \right) dy, \end{aligned}$$

with $\sigma_t^2 = \int_t^1 (1 - s)^2 ds = \frac{1}{3}(1 - t)^3$. Moreover,

$$\nabla u(t, x) = \frac{1}{\sqrt{2\pi\sigma_t}} \int_{\mathbb{R}} g(y) \frac{(y - x)}{\sigma_t^2} \exp \left(-\frac{(y - x)^2}{2\sigma_t^2} \right) dy. \quad (5.30)$$

By using the substitution $y = \sigma_t z$, we get

$$\begin{aligned} \nabla u(t, 0) &= \frac{1}{\sqrt{2\pi\sigma_t}} \int_{\mathbb{R}} g(\sigma_t z) z \exp \left(-\frac{z^2}{2} \right) dz \\ &= \frac{\sigma_t^{1/4}}{\sqrt{2\pi\sigma_t}} \int_{\mathbb{R}} \frac{1}{\sigma_t^{1/4}} \arctan \left(\sigma_t^{1/4} \frac{z}{|z|^{3/4}} \right) z \exp \left(-\frac{z^2}{2} \right) dy. \end{aligned}$$

By dominated convergence theorem we have

$$\int_{\mathbb{R}} \frac{1}{\sigma_t^{1/4}} \arctan \left(\sigma_t^{1/4} \frac{z}{|z|^{3/4}} \right) z \exp \left(-\frac{z^2}{2} \right) dy \xrightarrow{t \rightarrow 1^-} \int_{\mathbb{R}} |z|^{2-3/4} \exp \left(-\frac{z^2}{2} \right) dy > 0.$$

So,

$$\nabla u(t, 0)\sigma(t) \stackrel{t \rightarrow 1^-}{\sim} \frac{C(1-t)}{\sigma_t^{3/4}} = \frac{C}{(1-t)^{1/8}} \stackrel{t \rightarrow 1^-}{\rightarrow} +\infty.$$

Moreover, the equation (5.30) allows us to show that $(t, x) \mapsto \nabla u(t, x)$ is continuous on $[0, 1[\times \mathbb{R}$. So, for all $M > 0$, $\exists [t_1, t_2] \subset [0, 1[$ with $t_1 < t_2$, $\exists \varepsilon > 0$, such that $\forall x \in [-\varepsilon, \varepsilon]$, $\forall t \in [t_1, t_2]$ we have $|\nabla u(t, x)\sigma(t)| \geq M$. Since $Z_t = Z_t^{t, X_t^x} = \nabla u(t, X_t^x)\sigma(t)$, then

$$\mathbb{P}(|Z_{[t_1, t_2]}| \geq M) \geq \mathbb{P}(|X_{[t_1, t_2]}^x| \leq \varepsilon) > 0.$$

□

To construct this example, the idea is to stop the noise after the time $t = 1$ to obtain no random evolution in the time interval $[1, 2]$. In this example we need to have a function σ that depends on time. Let us see if it is possible to construct a counterexample when we are not in this situation. If we suppose that (HX0) holds, then the assumption (HX1'') is always true in dimension 1. So we will take $d = 2$ and

- $T = 2$,
- $\forall t \in [0, 2], \forall x = (x^1, x^2) \in \mathbb{R}^2, b(t, x) = \begin{pmatrix} 0 \\ h(t)x^1 \end{pmatrix}$,
- $\forall t \in [0, 2], \sigma(t) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$,
- $f = 0$

With these functions, (HX1) does not hold when $h \neq 0$. The solution of the SDE is

$$\begin{aligned} X_t^1 &= x^1 + W_t^1 \\ X_t^2 &= x^2 + x^1 \int_0^t h(s)ds + \int_0^t h(s)W_s^1 ds. \end{aligned}$$

As we try to obtain an explosion at time $t = 1$, we will take for h a continuous function such that $h|_{[0, 1[} > 0$ and $h|_{[1, 2]} = 0$. Moreover, we assume that $\forall x = (x^1, x^2) \in \mathbb{R}^2, g(x) = \tilde{g}(x^2)$ with $\tilde{g} : \mathbb{R} \rightarrow \mathbb{R}$ a bounded lower (or upper) semi-continuous function. As before, we denote $u(t, x) := Y_t^{t, x}$ with $(Y_t^{t, x}, Z_t^{t, x})$ the solution of the BSDE (5.5) where X verifies $X_t = x$. For $t \in [1, 2]$, obviously $u(t, x) = g(x) = \tilde{g}(x^2)$ and $Z_t^{t, x} = 0$. For $t \in [0, 1[$, we have

$$u(t, x) = \mathbb{E} \left[\tilde{g} \left(x^2 + x^1 \int_t^1 h(s)ds + \int_t^1 h(s)W_s^1 ds \right) \right].$$

Let us denote H a primitive of h . Then,

$$\begin{aligned} \int_t^1 h(s)W_s^1 ds &= H(1)W_1^1 - H(t)W_t^1 - \int_t^1 H(s)dW_s^1 \\ &= [H(1) - H(t)]W_t^1 + \int_t^1 [H(1) - H(s)]dW_s^1. \end{aligned}$$

So,

$$u(t, x) = \frac{1}{\sqrt{2\pi}\sigma_t} \int_{\mathbb{R}} \tilde{g}(y) \exp \left(-\frac{(y - x^2 - a_t x^1)^2}{2\sigma_t^2} \right) dy,$$

with $\sigma_t^2 = [H(1) - H(t)]^2 t + \int_t^1 [H(1) - H(s)]^2 ds$ and $a_t = [H(1) - H(t)]$. Moreover,

$$Z_t = \nabla u(t, X_t)\sigma = \frac{\partial u}{\partial x^1}(t, X_t),$$

and

$$\frac{\partial u}{\partial x^1}(t, x) = \frac{a_t}{\sqrt{2\pi}\sigma_t} \int_{\mathbb{R}} \tilde{g}(y) \frac{(y - x^2 - a_t x^1)}{\sigma_t^2} \exp\left(-\frac{(y - x^2 - a_t x^1)^2}{2\sigma_t^2}\right) dy.$$

So, By using the substitution $y - x^2 - a_t x^1 = \sigma_t z$, we get

$$\frac{\partial u}{\partial x^1}(t, x) = \frac{a_t}{\sqrt{2\pi}\sigma_t} \int_{\mathbb{R}} \tilde{g}(\sigma_t z + x^2 + a_t x^1) z \exp\left(-\frac{z^2}{2}\right) dz,$$

and

$$\left| \frac{\partial u}{\partial x^1}(t, x) \right| \leq C \frac{|a_t|}{\sigma_t} \int_{\mathbb{R}} |z| \exp\left(-\frac{z^2}{2}\right) dz \leq C$$

because $|a_t| \leq \sigma_t$. Finally, we obtain that Z is bounded on the interval $[0, 2]$ even if it is possible to show that $\frac{\partial u}{\partial x^2}(t, 0) \xrightarrow{t \rightarrow 1^-} +\infty$ for well-chosen functions \tilde{g} and h^4 . This example proves that (HX1'') is a not necessary but sufficient assumption. To be more precise, when σ does not depend on time we do not succeed to find an example of BSDE such that we have not the estimate

$$Z_t \leq C + C'(T - t)^{-1/2}.$$

So, the assumption (HX1'') seems to be artificially restrictive and unnecessary. But this question remains open.

5.6 Appendix

5.6.1 Proof of Lemma 5.10.

We have,

$$\prod_{i=0}^{2n-1} (1 + Mh_i) = \left(\prod_{i=0}^{n-1} (1 + Mh_i) \right) \left(\prod_{i=n}^{2n-1} (1 + Mh_i) \right).$$

Firstly,

$$\prod_{i=n}^{2n-1} (1 + Mh_i) \leq \left(1 + M \frac{T}{n} \right)^n \leq C.$$

Moreover, for $0 \leq i \leq n-1$,

$$h_i = t_{i+1} - t_i = T n^{-ai/n} (1 - e^{-\frac{a \ln n}{n}}) \leq T n^{-ai/n} a \frac{\ln n}{n},$$

thanks to the convexity of the exponential function. So

$$\begin{aligned} \prod_{i=0}^{n-1} (1 + Mh_i) &\leq \prod_{i=0}^{n-1} \left(1 + MT a n^{-ai/n} \frac{\ln n}{n} \right) \\ &= \exp \left(\sum_{i=0}^{n-1} \ln \left(1 + MT a n^{-ai/n} \frac{\ln n}{n} \right) \right) \\ &\leq \exp \left(\sum_{i=0}^{n-1} MT a (n^{-a/n})^i \frac{\ln n}{n} \right) \\ &\leq \exp \left(MT a \frac{\ln n}{n} \left(\frac{1 - (1/n^a)}{1 - (1/n^{(a/n)})} \right) \right) \\ &\leq \exp \left(MT a \frac{\ln n}{n} \frac{n^{a/n}}{n^{a/n} - 1} \right). \end{aligned}$$

But,

$$\frac{\ln n}{n} \frac{n^{a/n}}{n^{a/n} - 1} \sim \frac{\ln n}{n} \frac{1}{a \frac{\ln n}{n}} \sim \frac{1}{a},$$

when $n \rightarrow +\infty$. Thus, we have shown the result. \square

⁴Take for example $h(s) = (1 - s) \mathbb{1}_{s < 1}$ and $\tilde{g}(x) = \arctan\left(\frac{x}{|x|^{1/2}}\right)$.

5.6.2 Proof of Lemma 5.11.

Thanks to Lemma 5.10, we have

$$\frac{\prod_{i=0}^{n-1} \left(1 + M_1 h_i + M_2 \frac{h_i}{T-t_{i+1}}\right)}{\prod_{i=0}^{n-1} \left(1 + M_2 \frac{h_i}{T-t_{i+1}}\right)} = \prod_{i=0}^{n-1} \left(1 + \frac{M_1}{1 + M_2 \frac{h_i}{T-t_{i+1}}} h_i\right) \leq \prod_{i=0}^{n-1} (1 + M_1 h_i) \leq C.$$

So, we just have to show that

$$\prod_{i=0}^{n-1} \left(1 + M_2 \frac{h_i}{T-t_{i+1}}\right) \leq C n^{aM_2}.$$

But,

$$1 + M_2 \frac{h_i}{T-t_{i+1}} = 1 + M_2(n^{a/n} - 1).$$

So,

$$\begin{aligned} \prod_{i=0}^{n-1} \left(1 + M_2 \frac{h_i}{T-t_{i+1}}\right) &= \left(1 + M_2(n^{a/n} - 1)\right)^n \\ &= \exp\left(n \ln\left(1 + aM_2 \frac{\ln n}{n} + O\left(\frac{\ln^2 n}{n^2}\right)\right)\right) \\ &= \exp\left(aM_2 \ln n + O\left(\frac{\ln^2 n}{n}\right)\right) \sim n^{aM_2}, \end{aligned}$$

when $n \rightarrow +\infty$. Thus, we have shown the result. \square

5.6.3 Proof of Proposition 5.25.

We will prove this proposition as the authors of [36] do for their Proposition 4.2. In all the proof we omit the superscript N, ε for u, Y and Z to be more readable. Let $x_0, x'_0 \in \mathbb{R}^d$ and $t_0, t'_0 \in [0, T]$. By an argument of symmetry we are allowed to suppose that $t_0 \leq t'_0$. We have

$$|u(t_0, x_0) - u(t'_0, x'_0)| \leq |u(t_0, x_0) - u(t_0, x'_0)| + |u(t_0, x'_0) - u(t'_0, x'_0)|.$$

Let us begin with the first term. We will use a classical argument of linearization:

$$\begin{aligned} Y_t^{t_0, x_0} - Y_t^{t_0, x'_0} &= g_N(X_T^{t_0, x_0}) - g_N(X_T^{t_0, x'_0}) + \int_t^T \alpha_s \left(X_s^{t_0, x_0} - X_s^{t_0, x'_0}\right) + \beta_s \left(Y_s^{t_0, x_0} - Y_s^{t_0, x'_0}\right) ds \\ &\quad - \int_t^T \left(Z_s^{t_0, x_0} - Z_s^{t_0, x'_0}\right) d\tilde{W}_s, \end{aligned}$$

with

$$\begin{aligned} \alpha_s &= \begin{cases} \frac{f^\varepsilon(s, X_s^{t_0, x_0}, Y_s^{t_0, x'_0}, Z_s^{t_0, x'_0}) - f^\varepsilon(s, X_s^{t_0, x'_0}, Y_s^{t_0, x'_0}, Z_s^{t_0, x'_0})}{X_s^{t_0, x_0} - X_s^{t_0, x'_0}} & \text{if } X_s^{t_0, x_0} - X_s^{t_0, x'_0} \neq 0, \\ 0 & \text{elsewhere,} \end{cases} \\ \beta_s &= \begin{cases} \frac{f^\varepsilon(s, X_s^{t_0, x_0}, Y_s^{t_0, x_0}, Z_s^{t_0, x'_0}) - f^\varepsilon(s, X_s^{t_0, x_0}, Y_s^{t_0, x'_0}, Z_s^{t_0, x'_0})}{Y_s^{t_0, x_0} - Y_s^{t_0, x'_0}} & \text{if } Y_s^{t_0, x_0} - Y_s^{t_0, x'_0} \neq 0, \\ 0 & \text{elsewhere,} \end{cases} \\ \gamma_s &= \begin{cases} \frac{f^\varepsilon(s, X_s^{t_0, x_0}, Y_s^{t_0, x_0}, Z_s^{t_0, x_0}) - f^\varepsilon(s, X_s^{t_0, x_0}, Y_s^{t_0, x_0}, Z_s^{t_0, x'_0})}{|Z_s^{t_0, x_0} - Z_s^{t_0, x'_0}|^2} (Z_s^{t_0, x_0} - Z_s^{t_0, x'_0}) & \text{if } Z_s^{t_0, x_0} - Z_s^{t_0, x'_0} \neq 0, \\ 0 & \text{elsewhere,} \end{cases} \end{aligned}$$

and $d\tilde{W}_s := dW_s - \gamma_s ds$. By a BMO argument, there exists a probability \mathbb{Q} under which \tilde{W} is a Brownian motion. Then we apply a classical transformation to obtain

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} \left[e^{\int_{t_0}^t \beta_s ds} \left(Y_t^{t_0, x_0} - Y_t^{t_0, x'_0} \right) \right] &= \mathbb{E}^{\mathbb{Q}} \left[e^{\int_{t_0}^T \beta_s ds} \left(g_N(X_T^{t_0, x_0}) - g_N(X_T^{t_0, x'_0}) \right) \right. \\ &\quad \left. + \int_{t_0}^T \alpha_s e^{\int_{t_0}^s \beta_u du} \left(X_s^{t_0, x_0} - X_s^{t_0, x'_0} \right) ds \right], \end{aligned}$$

and

$$|u(t_0, x_0) - u(t_0, x'_0)| \leq C \left(\mathbb{E}^{\mathbb{Q}} \left[\omega \left(\left| X_T^{t_0, x_0} - X_T^{t_0, x'_0} \right| \right) \right] + \int_{t_0}^T \mathbb{E}^{\mathbb{Q}} \left[\left| X_s^{t_0, x_0} - X_s^{t_0, x'_0} \right| \right] ds \right),$$

with ω a modulus of continuity of g that is also a modulus of continuity for g_N . We are allowed to suppose that ω is concave⁵, so Jensen's inequality gives us

$$|u(t_0, x_0) - u(t_0, x'_0)| \leq C \left(\omega \left(\mathbb{E}^{\mathbb{Q}} \left[\left| X_T^{t_0, x_0} - X_T^{t_0, x'_0} \right| \right] \right) + \int_{t_0}^T \mathbb{E}^{\mathbb{Q}} \left[\left| X_s^{t_0, x_0} - X_s^{t_0, x'_0} \right| \right] ds \right).$$

By using the fact that b is bounded we can prove the following proposition exactly as authors of [36] do for their Proposition 4.7:

Proposition 5.29 $\exists C > 0$ that does not depend on N and ε such that $\forall t, t' \in [0, T], \forall x, x' \in \mathbb{R}^d, \forall s \in [0, T]$,

$$\mathbb{E}^{\mathbb{Q}} \left[\left| X_s^{t, x} - X_s^{t', x'} \right| \right] \leq C \left(|x - x'| + |t - t'|^{1/2} \right).$$

Then,

$$|u(t_0, x_0) - u(t_0, x'_0)| \leq C \left(\omega(|x_0 - x'_0|) + |x_0 - x'_0| \right).$$

Now we will study the second term:

$$|u(t_0, x'_0) - u(t'_0, x'_0)| = \left| Y_{t_0}^{t_0, x'_0} - Y_{t'_0}^{t'_0, x'_0} \right| \leq \left| Y_{t_0}^{t_0, x'_0} - Y_{t'_0}^{t_0, x'_0} \right| + \left| Y_{t'_0}^{t_0, x'_0} - Y_{t'_0}^{t'_0, x'_0} \right|.$$

Firstly,

$$\left| Y_{t'_0}^{t_0, x'_0} - Y_{t'_0}^{t'_0, x'_0} \right| \leq \left| \int_{t_0}^{t'_0} f(s, x'_0, Y_s^{t_0, x'_0}, 0) ds \right| \leq C |t_0 - t'_0|.$$

Moreover, as for the first term we have

$$\begin{aligned} \mathbb{E}^{\mathbb{Q}} \left[e^{\int_{t_0}^t \beta_s ds} \left(Y_t^{t_0, x'_0} - Y_t^{t'_0, x'_0} \right) \right] &= \mathbb{E}^{\mathbb{Q}} \left[e^{\int_{t_0}^T \beta_s ds} \left(g_N(X_T^{t_0, x'_0}) - g_N(X_T^{t'_0, x'_0}) \right) \right. \\ &\quad \left. + \int_{t_0}^T \alpha_s e^{\int_{t_0}^s \beta_u du} \left(X_s^{t_0, x'_0} - X_s^{t'_0, x'_0} \right) ds \right], \end{aligned}$$

and

$$\left| Y_{t_0}^{t_0, x'_0} - Y_{t'_0}^{t'_0, x'_0} \right| \leq C \left(\omega(|t_0 - t'_0|^{1/2}) + |t_0 - t'_0|^{1/2} \right).$$

Finally,

$$|u(t_0, x'_0) - u(t'_0, x'_0)| \leq C \left(\omega(|t_0 - t'_0|^{1/2}) + |t_0 - t'_0|^{1/2} \right),$$

and

$$|u(t_0, x_0) - u(t'_0, x'_0)| \leq C \left(\omega(|x_0 - x'_0|) + \omega(|t_0 - t'_0|^{1/2}) + |x_0 - x'_0| + |t_0 - t'_0|^{1/2} \right).$$

So u is uniformly continuous on $[0, T] \times \mathbb{R}^d$ and this function has a modulus of continuity that does not depend on N and ε . Moreover, we are allowed to suppose that this modulus of continuity is concave. \square

⁵There exist two positive constants a and b such that $\omega(x) \leq ax + b$. Then the concave hull of $x \mapsto \omega(x) \vee (\mathbb{1}_{x \geq 1}(ax + b))$ is also a modulus of continuity of g .

Chapitre 6

Résultats numériques

Ce chapitre a pour but d'illustrer les résultats théoriques obtenus dans le chapitre précédent. Plus précisément, nous nous intéressons à la simulation numérique d'EDSRs de la forme

$$Y_t = g(X_T) + \int_t^T f(s, X_s, Y_s, Z_s) ds - \int_t^T Z_s dW_s,$$

avec X le processus solution de l'EDS

$$X_t = x + \int_0^t b(s, X_s) ds + \int_0^t \sigma(s) dW_s.$$

b, σ, g et f vérifient les hypothèses suivantes :

- $b : [0, T] \times \mathbb{R}^d \rightarrow \mathbb{R}^d$ est une fonction mesurable telle que $\exists M_b, K_b \in \mathbb{R}^+, \forall t, t' \in [0, T], \forall x, x' \in \mathbb{R}^d$,

$$\begin{aligned} |b(t, x)| &\leq M_b, \\ |b(t, x) - b(t', x')| &\leq K_b(|x - x'| + |t - t'|^{1/2}), \end{aligned}$$

- $\sigma : [0, T] \rightarrow \mathbb{R}^{d \times d}$ est une fonction mesurable telle que $\exists M_\sigma, K_\sigma \in \mathbb{R}^+, \forall t, t' \in [0, T], \forall x, x' \in \mathbb{R}^d$,

$$\begin{aligned} |\sigma(t)| &\leq M_\sigma, \\ |\sigma(t) - \sigma(t')| &\leq K_\sigma |t - t'|, \end{aligned}$$

- $g : \mathbb{R}^d \rightarrow \mathbb{R}$ est une fonction mesurable telle que $\exists M_g \in \mathbb{R}^+, \exists \alpha \in]0, 1]$,

$$\forall x \in \mathbb{R}^d, \quad |g(x)| \leq M_g, \quad \sup_{x, x' \in \mathbb{R}^d} \frac{|g(x) - g(x')|}{|x - x'|^\alpha} < +\infty,$$

- $f : [0, T] \times \mathbb{R}^d \times \mathbb{R} \times \mathbb{R}^{1 \times d} \rightarrow \mathbb{R}$ est une fonction mesurable telle que $\exists M_f, K_{f,t}, K_{f,x}, K_{f,y}, K_{f,z} \in \mathbb{R}^+, \forall t, t' \in [0, T], \forall x, x' \in \mathbb{R}^d, \forall y, y' \in \mathbb{R}, \forall z, z' \in \mathbb{R}^{1 \times d}$,

$$\begin{aligned} |f(t, x, y, z)| &\leq M_f(1 + |y| + |z|^2), \\ |f(t, x, y, z) - f(t', x', y', z')| &\leq K_{f,t}|t - t'|^{1/2} + K_{f,x}|x - x'| + K_{f,y}|y - y'| \\ &\quad + (K_{f,z} + L_{f,z}(|z| + |z'|))|z - z'|. \end{aligned}$$

De plus, nous supposons qu'il existe $M_1, M_2 \in \mathbb{R}^+$ telles que, \mathbb{P} -p.s.,

$$|Z_t| \leq M_1 + \frac{M_2}{(T - t)^{1/2}}, \quad \forall t \in [0, T[.$$

Cette majoration est obtenue dans les théorèmes 5.4 et 5.27 modulo des hypothèses supplémentaires sur b et σ .

6.1 Description de l'algorithme

Nous allons, dans une première partie, rappeler le schéma de discrétisation temporelle étudié dans le chapitre 5 puis, dans une seconde partie, nous détaillerons la méthode employée pour estimer les espérances conditionnelles.

6.1.1 Discrétisation temporelle

Commençons par donner quelques notations. Nous noterons n le nombre de pas de temps et $0 = t_0 < t_1 < \dots < t_n = T$ la grille des temps de discrétisation. Dans toute la suite K désignera une constante positive telle que

$$K > 4L_{f,z}^2 M_2^2.$$

On pose $\varepsilon := Tn^{-a}$ et $N := n^b$ avec

$$b := \frac{1-\alpha}{(2-\alpha)(2+K)-2+2\alpha} \quad \text{et} \quad a := \frac{1+2b}{2+K}.$$

On définit également la constante c par la formule

$$c := 1 + \frac{3\alpha - 4}{(2-\alpha)(2+K)-2+2\alpha}.$$

Enfin la grille de discrétisation est donnée par

$$\begin{cases} t_k &= T \left(1 - \left(\frac{\varepsilon}{T} \right)^{k/n_1} \right), & 0 \leq k \leq n_1, \\ t_k &= T - \left(\frac{n-k}{n-n_1} \right) \varepsilon, & n_1 \leq k \leq n, \end{cases}$$

avec n_1 définie ainsi :

- si $K \leq \frac{2-3\alpha}{2-\alpha}$ alors $n_1 = n - 1$,
- si $K > \frac{2-3\alpha}{2-\alpha}$ alors n_1 est le plus grand entier vérifiant $n_1 + \lceil n_1^c \rceil \leq n$.

L'EDS est remplacée par le schéma d'Euler classique :

$$\begin{cases} X_0^n &= x \\ X_{t_{k+1}}^n &= X_{t_k}^n + (t_{k+1} - t_k)b(t_k, X_{t_k}^n) + \sigma(t_k)(W_{t_{k+1}} - W_{t_k}), \quad 0 \leq k \leq n. \end{cases}$$

Afin d'alléger les notations du chapitre précédent, nous noterons (Y^n, Z^n) la solution de l'EDSR discrétisée en temps en lieu et place de $(Y^{N,\varepsilon,n}, Z^{N,\varepsilon,n})$. (Y^n, Z^n) est définie par une équation de programmation dynamique :

$$\begin{cases} Y_{t_n}^n &= g_N(X_{t_n}^n) \\ Z_{t_k}^n &= \rho_{t_{k+1}} \left(\frac{1}{t_{k+1} - t_k} \mathbb{E} \left[Y_{t_{k+1}}^n (W_{t_{k+1}} - W_{t_k}) \middle| \mathcal{F}_{t_k} \right] \right), \quad 0 \leq k \leq n-1, \\ Y_{t_k}^n &= \mathbb{E} \left[Y_{t_{k+1}}^n + (t_{k+1} - t_k) f^\varepsilon(t_k, X_{t_k}^n, Y_{t_{k+1}}^n, Z_{t_k}^n) \middle| \mathcal{F}_{t_k} \right], \quad 0 \leq k \leq n-1, \end{cases}$$

avec $\rho_s : \mathbb{R}^{1 \times d} \rightarrow \mathbb{R}^{1 \times d}$ la projection sur la boule fermée

$$B \left(0, M_{z,1} + \frac{M_{z,2}}{(T-s)^{1/2}} \right),$$

f^ε une modification de f définie par

$$f^\varepsilon(s, x, y, z) := \mathbb{1}_{s \leq T-\varepsilon} f(s, x, y, z) + \mathbb{1}_{s > T-\varepsilon} f(s, x, y, 0),$$

et g_N une approximation N -lipschitz de g donnée par

$$g_N(x) = \inf \{ g(u) + N |x - u| \mid u \in \mathbb{R}^d \}.$$

6.1.2 Discrétisation spatiale

Pour pouvoir obtenir un schéma implantable sur un ordinateur il reste à estimer les espérances conditionnelles. Pour cela nous allons utiliser le caractère markovien de X^n qui permet de montrer facilement que Y^n et Z^n sont des fonctions déterministes de X^n .

Proposition 6.1 *Supposons vérifiées les hypothèses décrites au début de ce chapitre sur f , g , b et σ . Alors on a*

$$Y_{t_k}^n = y_k^n(X_{t_k}^n), \quad Z_{i,t_k}^n = z_{i,k}^n(X_{t_k}^n), \quad 0 \leq k \leq n, \quad 1 \leq i \leq d,$$

où $(y_k^n(\cdot))_k, (z_{i,k}^n(\cdot))_{i,k}$ sont des fonctions mesurables.

L'idée développée dans les articles [48, 49] consiste à approcher y_k^n et $(z_{i,k}^n)_i$ par leur projection sur des sous-espaces vectoriels de dimension finie engendrés par les bases p_k et $(q_{i,k})_i$. On définit ainsi de nouvelles fonctions :

$$y_k^{n,\delta} := \text{Proj}_{\text{Vect}\{p_k\}} y_k^n = \alpha_k p_k, \quad z_{i,k}^{n,\delta} := \text{Proj}_{\text{Vect}\{q_{i,k}\}} z_{i,k}^n = \beta_{i,k} q_k, \quad 0 \leq k \leq n, \quad 1 \leq i \leq d,$$

où α_k et $\beta_{i,k}$ sont des ensembles de coefficients¹. Enfin, ces coefficients $(\alpha_k)_k$ et $(\beta_{i,k})_{i,k}$ sont eux-mêmes estimés à l'aide d'une méthode de Monte-Carlo. Pour ce faire, il convient de réaliser M tirages indépendants de X^n : nous noterons $X^{n,m}$ le m ème tirage et nous poserons $\Delta W_{i,k}^m := W_{i,t_{k+1}}^m - W_{i,t_k}^m$. L'algorithme final se présente toujours sous forme rétrograde :

- Initialisation pour $k = n$ par $y_n^{n,\delta,M} = g_N$.
- À l'instant de discrétisation t_k , on calcule les coefficients $(\beta_{i,k}^M)_i$ par une méthode des moindres carrés :

$$\beta_{i,k}^M := \arg \min_{\beta} \frac{1}{M} \sum_{m=1}^M \left| y_{k+1}^{n,\delta,M}(X_{t_{k+1}}^{n,m}) \frac{\Delta W_{i,k}^m}{h_k} - \beta q_{i,k}(X_{t_k}^{n,m}) \right|^2.$$

On pose alors

$$z_k^{n,\delta,M} = (z_{i,k}^{n,\delta,M})_i := \rho_{t_{k+1}} \left((\beta_{i,k}^M q_{i,k})_i \right).$$

- Toujours à l'instant de discrétisation t_k , on calcule ensuite les coefficients α_k^M par une méthode similaire :

$$\alpha_k^M := \arg \min_{\alpha} \frac{1}{M} \sum_{m=1}^M \left| y_{k+1}^{n,\delta,M}(X_{t_{k+1}}^{n,m}) + h_k f^\varepsilon \left(t_k, X_{t_k}^{n,m}, y_{k+1}^{n,\delta,M}(X_{t_{k+1}}^{n,m}), z_k^{n,\delta,M}(X_{t_k}^{n,m}) \right) - \alpha p_k(X_{t_k}^{n,m}) \right|^2.$$

On pose alors

$$y_k^{n,\delta,M} := \alpha_k^M p_k.$$

- On itère ces calculs jusqu'à l'instant t_0 .

La solution $(Y_t, Z_t)_{t \in [0,T]}$ de notre EDSR initiale est alors approchée numériquement par

$$Y_{t_k}^{n,\delta,M} := y_k^{n,\delta,M}(X_{t_k}^n), \quad Z_{i,t_k}^{n,\delta,M} := z_{i,k}^{n,\delta,M}(X_{t_k}^n), \quad 0 \leq k \leq n.$$

Il reste maintenant à définir les bases de fonctions $(p_k)_k$ et $(q_{i,k})_{i,k}$. Pour simplifier, nous ne les ferons pas dépendre de i et nous ne ferons pas de distinction entre p et q : nous noterons p_k cette base au temps t_k . Pour définir ces dernières, nous allons découper \mathbb{R}^d en petits hypercubes d'arête δ :

$$\mathbb{R}^d = \bigcup_{(i_1, \dots, i_d) \in \mathbb{Z}^d} D_{i_1, \dots, i_d}, \quad \text{avec} \quad D_{i_1, \dots, i_d} :=]i_1\delta, (i_1+1)\delta] \times \dots \times]i_d\delta, (i_d+1)\delta].$$

Nous pourrions alors définir p_k comme la famille des fonctions indicatrices associées à l'ensemble des hypercubes $(\mathbb{1}_{D_{i_1, \dots, i_d}}(\cdot))_{(i_1, \dots, i_d) \in \mathbb{Z}^d}$, mais cette base n'est pas de cardinal fini. Néanmoins, nous pouvons nous contenter des hypercubes dans lesquels « tombent » les $(X_{t_k}^{n,m})_{1 \leq m \leq M}$:

$$p_k(\cdot) := \left(\mathbb{1}_{D_{i_1, \dots, i_d}}(\cdot) \right)_{(i_1, \dots, i_d) \in \mathcal{E}_k},$$

¹ Afin de ne pas alourdir inutilement les notations nous n'avons pas indexé les bases de fonctions et les coefficients correspondants.

avec

$$\mathcal{E}_k := \{(i_1, \dots, i_d) \in \mathbb{Z}^d \mid \exists m \in \{1, \dots, M\}, X_{t_k}^{n,m} \in D_{i_1, \dots, i_d}\}.$$

Nous utiliserons principalement ces bases par la suite. De plus nous pourrions éventuellement considérer une variante de ces bases qui consiste à prendre sur chaque hypercube une base de polynômes locaux de degré 1. On a alors

$$p_k(\cdot) := \left(\mathbb{1}_{D_{i_1, \dots, i_d}}(\cdot), x_1 \mathbb{1}_{D_{i_1, \dots, i_d}}(\cdot), \dots, x_d \mathbb{1}_{D_{i_1, \dots, i_d}}(\cdot) \right)_{(i_1, \dots, i_d) \in \mathcal{E}_k}.$$

D'autres choix pour ces bases sont également envisageables. Nous ne les utiliserons pas ici mais nous pouvons tout de même citer quelques exemples :

- On peut prendre sur chaque hypercube une base de polynômes locaux de degré supérieur à 1.
- Il est possible de considérer une base de polynômes globaux sur \mathbb{R}^d .
- On peut utiliser les fonctions indicatrices des pavages de Voronoï associés à R simulations indépendantes de X^n .

Une étude numérique approfondie comparant ces différentes possibilités est disponible dans la thèse [63].

6.1.3 Choix des paramètres

Au final, l'algorithme dépend de trois paramètres. Dans le cas des EDSRs lipschitz, la thèse [63] propose une étude théorique et numérique complète de l'erreur globale en fonction de ceux-ci. Il en résulte que le comportement de l'algorithme est très sensible à ces derniers et qu'il convient de les faire varier correctement afin d'obtenir une convergence. Par exemple, si l'on se contente d'augmenter n en fixant M et δ , l'erreur globale peut exploser. Afin de pallier ce risque, Lemor propose de fixer n , δ et M ainsi :

$$n = \lfloor n_0(\sqrt{2})^{j-1} \rfloor, \quad M = \lfloor M_0(\sqrt{2})^{(j-1)\alpha_M} \rfloor, \quad \delta = \frac{\delta_0}{(\sqrt{2})^{(j-1)\alpha_\delta}}, \quad j \in \mathbb{N}^*.$$

Nous prendrons par la suite $n_0 = 2$, $M_0 = 50$ et $\delta_0 = 4$. Il reste alors à fixer α_M et α_δ suffisamment grands : en effet, théoriquement il convient d'avoir $\alpha_\delta > 0.5$ et pour α_δ fixé, il existe un seuil pour α_M en dessous duquel l'algorithme ne converge pas et au dessus duquel il converge. Au vu des résultats numériques de Lemor, nous prendrons $\alpha_\delta = 1, 5$ et $\alpha_M \in \{2, 5; 3\}$.

| j | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|-------------------------|----|-------|-------|-------|------|--------|--------|--------|--------|
| n | 2 | 2 | 4 | 5 | 8 | 11 | 16 | 22 | 32 |
| M ($\alpha_M = 2, 5$) | 50 | 118 | 282 | 672 | 1600 | 3805 | 9050 | 21526 | 51200 |
| M ($\alpha_M = 3$) | 50 | 141 | 400 | 1131 | 3200 | 9050 | 25600 | 72407 | 204800 |
| δ | 4 | 2,378 | 1,414 | 0,841 | 0,5 | 0,2973 | 0,1768 | 0,1051 | 0,0625 |

6.2 Convergence de l'algorithme

Tous les résultats numériques de cette partie sont obtenus en lançant 30 fois l'algorithme.

6.2.1 Exemples considérés

Nous allons nous placer en dimension $d = 1$ et considérer pour X un simple mouvement brownien :

$$X_t = W_t, \quad 0 \leq t \leq T.$$

Les fonctions f et g sont données par :

$$f(\cdot, \cdot, \cdot, z) = -\frac{|z|^2}{2}, \quad z \in \mathbb{R},$$

$$g(x) = x^\alpha \mathbb{1}_{0 \leq x \leq 1} + \mathbb{1}_{x > 1}, \quad x \in \mathbb{R},$$

avec $\alpha \in]0, 1]$. Pour l'approximation lipschitz de g nous posons

$$g_N(x) = \begin{cases} g(x) & \text{si } x \notin \left[0, N^{\frac{-1}{1-\alpha}}\right], \\ Nx & \text{sinon.} \end{cases}$$

Une transformation exponentielle classique permet de résoudre de manière exacte cette EDSR quadratique :

$$Y_t = \ln \mathbb{E} \left[e^{g(W_T)} | \mathcal{F}_t \right].$$

Ainsi, on a

$$Y_0 = \ln \mathbb{E} \left[e^{g(W_T)} \right] = \ln(u(0)),$$

avec

$$u(x) = \frac{1}{\sqrt{2\pi}} \int_{\mathbb{R}} e^{g(y)} e^{-\frac{(y-x)^2}{2}} dy,$$

ce qui nous donne finalement

$$Y_0 = \ln \left(\frac{1}{2} + e\mathbb{P}(W_1 \geq 1) + \frac{1}{\sqrt{2\pi}} \int_0^1 e^{y^\alpha} e^{-y^2/2} dy \right).$$

De plus, on a

$$Z_0 = \frac{u'(0)}{u(0)},$$

avec

$$u'(0) = \frac{e^{1/2} - 1}{\sqrt{2\pi}} + \frac{1}{\sqrt{2\pi}} \int_0^1 y e^{y^\alpha} e^{-y^2/2} dy.$$

Voici les valeurs obtenues pour différents α :

| α | Y_0 | Z_0 |
|----------|--------|--------|
| 1 | 0,4015 | 0,3767 |
| 0,5 | 0,4659 | 0,3797 |
| 0,1 | 0,5729 | 0,3743 |
| 0 | 0,6201 | 0,3687 |

Il convient également de calculer explicitement l'estimation temporelle du processus Z provenant du théorème 5.4 en reprenant sa démonstration. Supposons que g est différentiable, alors

$$\nabla Y_t = \nabla g(X_T) \nabla X_T - \int_t^T \nabla Z_s d\tilde{W}_s$$

avec $d\tilde{W}_s = dW_s - Z_s ds$ qui est un mouvement brownien sous la probabilité \mathbb{Q} . On a $\nabla X_t = \text{Id}$ et $Z_t = \nabla Y_t$, donc Z est une \mathbb{Q} -martingale et $|Z|^2$ est une \mathbb{Q} -sous-martingale :

$$|Z_t|^2 (T - t) \leq \mathbb{E}^{\mathbb{Q}} \left[\int_t^T |Z_s|^2 ds \middle| \mathcal{F}_t \right],$$

cette inégalité restant vraie lorsque g est uniquement semi-continue inférieurement ou supérieurement ce qui est notre cas. Or,

$$\begin{aligned} Y_t &= g(X_T) + \int_t^T \frac{|Z_s|^2}{2} ds - \int_t^T Z_s dW_s \\ &= g(X_T) - \int_t^T \frac{|Z_s|^2}{2} ds - \int_t^T Z_s d\tilde{W}_s. \end{aligned}$$

Finalement,

$$\frac{1}{2} \mathbb{E}^{\mathbb{Q}} \left[\int_t^T |Z_s|^2 ds \middle| \mathcal{F}_t \right] = \mathbb{E}^{\mathbb{Q}} [Y_T - Y_t | \mathcal{F}_t].$$

Une estimation classique sur Y nous donne la majoration $|Y| \leq 1$ ce qui nous permet d'obtenir une première estimation pour Z :

$$|Z_t| \leq \frac{2}{\sqrt{T-t}}, \quad t \in [0, T[. \quad (6.1)$$

Comme nous l'avons vu dans la démonstration du théorème 5.23, il est également possible de développer le terme $\mathbb{E}^{\mathbb{Q}} [Y_T - Y_t | \mathcal{F}_t]$:

$$|\mathbb{E}^{\mathbb{Q}} [Y_T - Y_t | \mathcal{F}_t]| \leq \mathbb{E}^{\mathbb{Q}} \left[\left| u(T, \tilde{W}_T) - u(t, \tilde{W}_t) \right| \middle| \mathcal{F}_t \right],$$

avec

$$u(t, x) = -\ln \mathbb{E}^{\mathbb{Q}} \left[e^{-g(\tilde{W}_{T-t} + x)} \right].$$

Or

$$\begin{aligned} |u(t, x) - u(t', x')| &\leq e \left| \mathbb{E}^{\mathbb{Q}} \left[e^{-g(\tilde{W}_{T-t} + x)} - e^{-g(\tilde{W}_{T-t'} + x')} \right] \right| \\ &\leq e \mathbb{E}^{\mathbb{Q}} \left| g(\tilde{W}_{T-t} + x) - g(\tilde{W}_{T-t'} + x') \right| \\ &\leq e \left(\mathbb{E} \left[\left| \tilde{W}_{T-t} - \tilde{W}_{T-t'} \right|^\alpha \right] + |x - x'|^\alpha \right) \\ &\leq e \left(|t - t'|^{\alpha/2} + |x - x'|^\alpha \right), \end{aligned}$$

ce qui nous donne

$$\begin{aligned} |\mathbb{E}^{\mathbb{Q}} [Y_T - Y_t | \mathcal{F}_t]| &\leq e \left(|T - t|^{\alpha/2} + \mathbb{E}^{\mathbb{Q}} \left[\left| \tilde{W}_T - \tilde{W}_t \right|^\alpha \middle| \mathcal{F}_t \right] \right) \\ &\leq 2e |T - t|^{\alpha/2}. \end{aligned}$$

Nous obtenons ainsi une seconde estimation sur Z plus fine :

$$|Z_t| \leq \frac{2\sqrt{e}}{|T - t|^{\frac{2-\alpha}{4}}}, \quad t \in [0, T[. \quad (6.2)$$

Enfin, rappelons que comme g est approchée par g_N qui est une fonction N -lipschitz, alors le théorème 5.3 nous assure que Z^n est borné :

$$|Z^n| \leq N = n^{-\frac{1-\alpha}{(2-\alpha)(2+K)-2+2\alpha}}. \quad (6.3)$$

6.2.2 Utilité de la projection pour Z

Dans un premier temps, nous allons traiter de l'utilité de la projection. Pour cela nous considérons le schéma suivant :

- pas de projection pour Z ,
- grille de discrétisation temporelle uniforme,
- approximation lipschitz de g .

Voici les résultats obtenus pour les différents α_M :

- $\alpha_M = 2, 5$

| j | moyenne Y_0 | écart type Y_0 | moyenne Z_0 | écart type Z_0 |
|-----|----------------|------------------|---------------|------------------|
| 1 | 0,3872 | 0,09461 | 0,2127 | 0,1363 |
| 2 | 0,3950 | 0,05454 | 0,1905 | 0,05892 |
| 3 | 0,4105 | 0,03442 | 0,1463 | 0,05892 |
| 4 | 0,4259 | 0,02999 | 0,2174 | 0,06622 |
| 5 | $4,017 * 10^7$ | $2,200 * 10^8$ | 4910 | $2,689 * 10$ |
| 6 | ∞^2 | ∞ | ∞ | ∞ |

- $\alpha_M = 3$

| j | moyenne Y_0 | écart type Y_0 | moyenne Z_0 | écart type Z_0 |
|-----|--------------------|-------------------|-------------------|-------------------|
| 1 | 0,3841 | 0,08744 | 0,1945 | 0,1151 |
| 2 | 0,3941 | 0,05596 | 0,1926 | 0,06994 |
| 3 | 0,4058 | 0,05054 | 0,1607 | 0,1194 |
| 4 | 0,4146 | 0,01736 | 0,1957 | 0,03452 |
| 5 | $1,401 * 10^{16}$ | $7,671 * 10^{16}$ | $9,274 * 10^8$ | $5,019 * 10^8$ |
| 6 | $2,469 * 10^{169}$ | ∞ | $4,444 * 10^{84}$ | $2,434 * 10^{85}$ |
| 7 | ∞ | ∞ | ∞ | ∞ |

²L'infini numérique correspond à $1,7 * 10^{308}$.

Ces résultats prouvent qu'il est nécessaire d'utiliser une projection pour Z afin d'assurer une non explosion de la solution. Ainsi, nous allons modifier la projection pour éviter que $Z_{t_{n-1}}^{n,\delta,M}$ soit non borné. Dorénavant nous poserons

$$z_k^{n,\delta,M} = \left(z_{i,k}^{n,\delta,M} \right)_i := \rho_{t_k} \left((\beta_{i,k}^M q_{i,k})_i \right).$$

6.2.3 Utilisation de l'algorithme non optimisé

Dans cette partie, nous allons utiliser le schéma suivant :

- projection pour Z en utilisant la borne

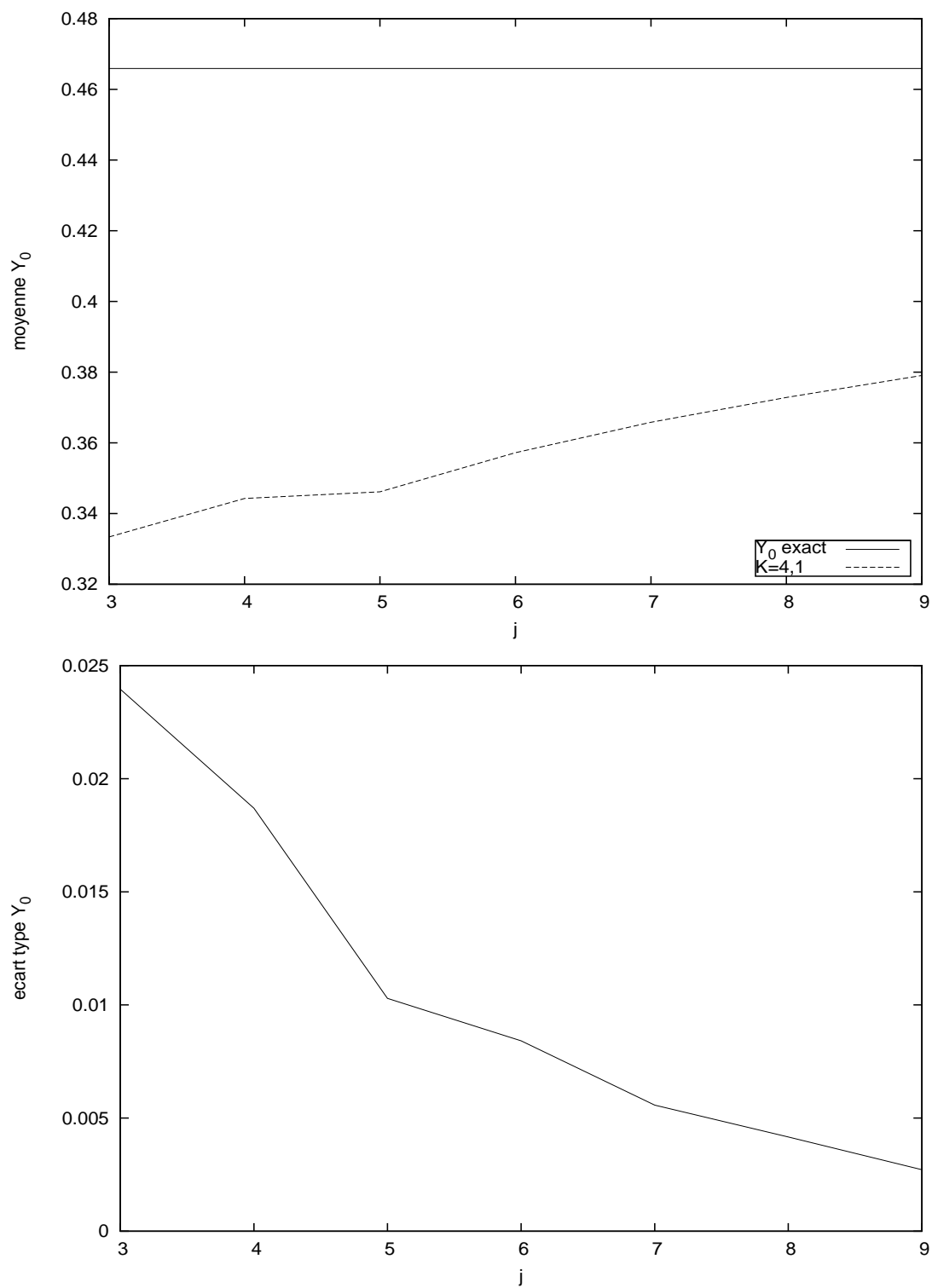
$$|Z_t| \leq \frac{2}{\sqrt{T-t}},$$

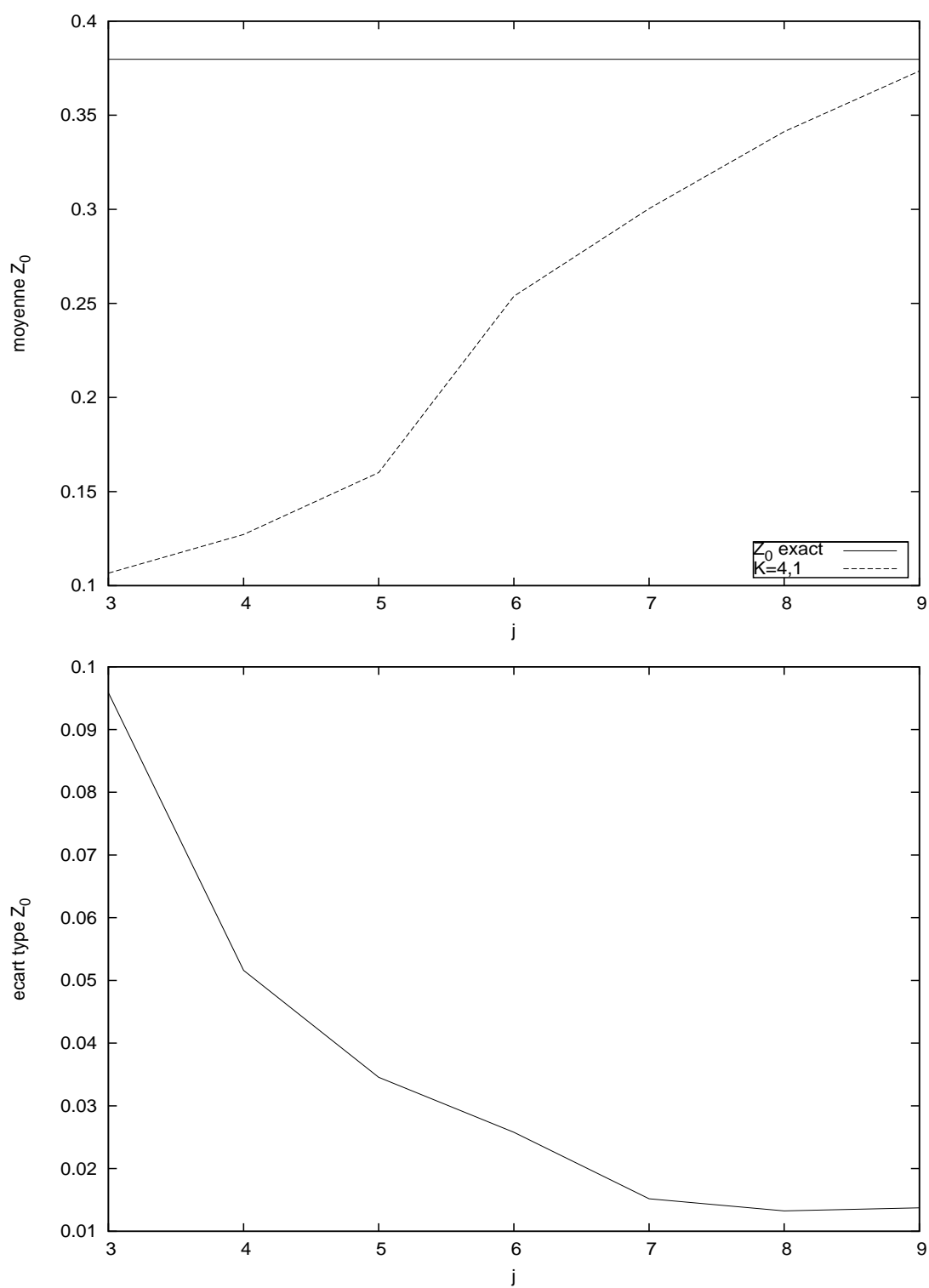
- grille de discrétisation temporelle non uniforme avec $K = 4, 1$,
- approximation lipschitz de g .

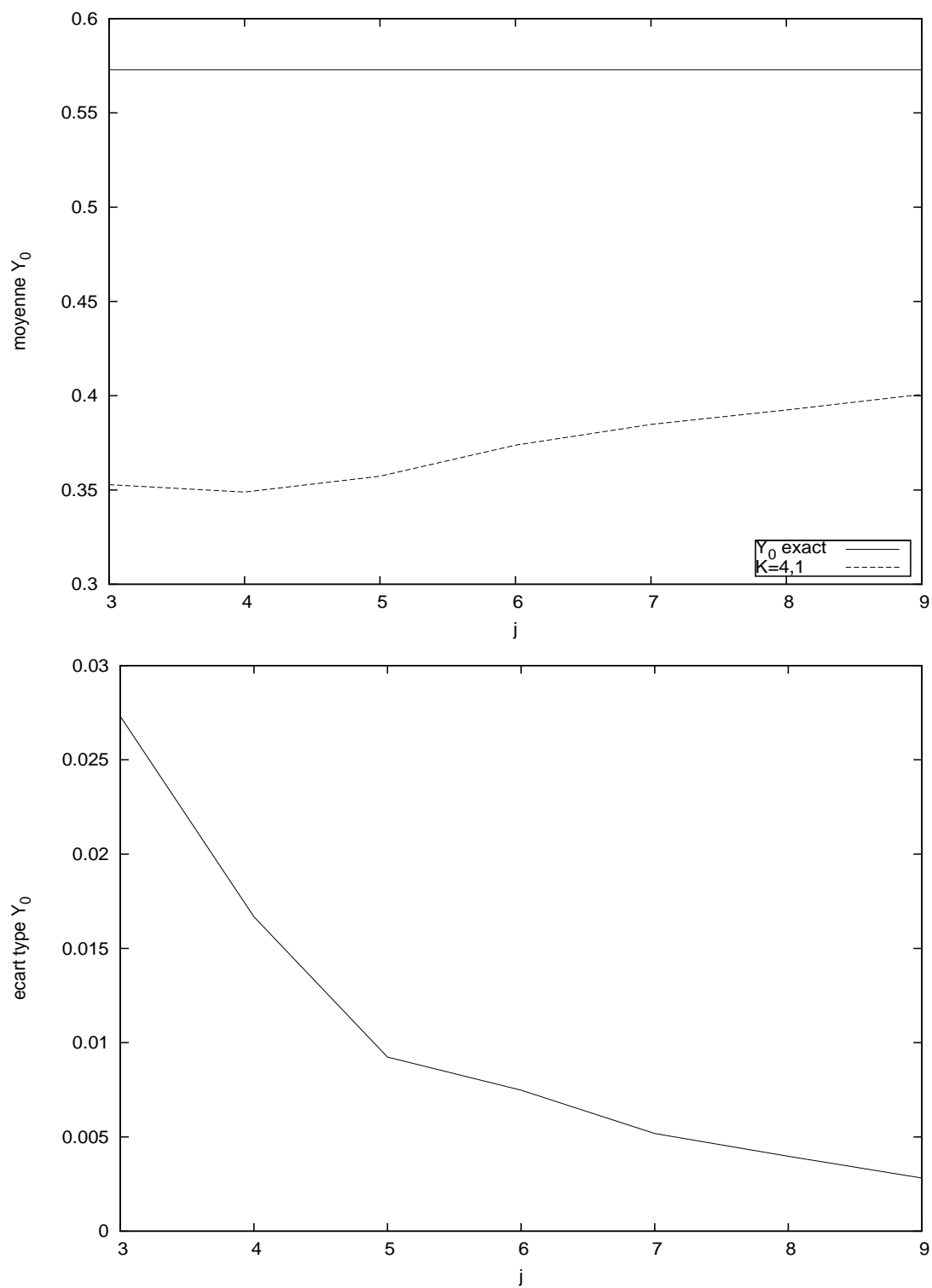
Comme $K > \frac{2-3\alpha}{2-\alpha}$, la grille de discrétisation est formée de deux parties aux comportements distincts : $t_0 < \dots < t_{n_1}$ et $t_{n_1} < \dots < t_n$. Les valeurs de n_1 , pour $\alpha = 0, 5$ et $\alpha = 0, 1$, sont les suivantes :

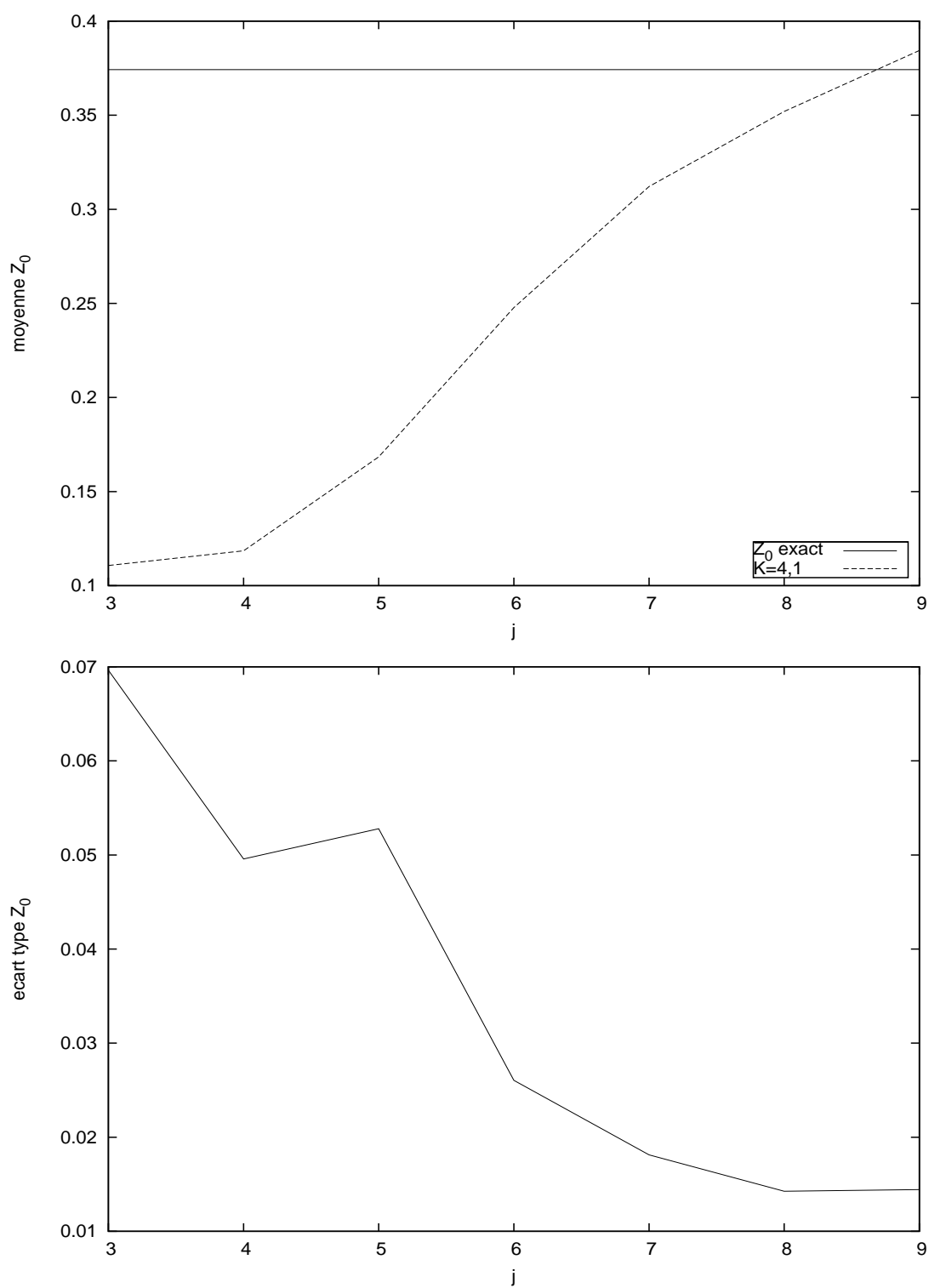
| j | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|-----------------------|---|---|---|---|---|----|----|----|----|
| n | 2 | 2 | 4 | 5 | 8 | 11 | 16 | 22 | 32 |
| $n_1 (\alpha = 0, 5)$ | 1 | 1 | 2 | 2 | 4 | 7 | 10 | 15 | 23 |
| $n_1 (\alpha = 0, 1)$ | 1 | 1 | 2 | 3 | 5 | 7 | 11 | 16 | 24 |

Les figures 6.1 et 6.2 donnent les résultats de la simulation pour $\alpha = 0, 5$ tandis que les figures 6.3 et 6.4 concernent le cas $\alpha = 0, 1$. Nous pouvons constater que la vitesse de convergence de Y_0 est extrêmement lente. Par contre, il ne semble pas possible de conclure quant à la convergence de Z_0 : il conviendrait de faire d'autres simulations pour j plus grand. Cette vitesse faible fait que l'algorithme n'est pas exploitable en pratique : nous allons par la suite utiliser une grille de discrétisation temporelle non uniforme avec $K = 0, 1$.

FIG. 6.1 – Convergence de Y_0 pour une grille de discrétisation non uniforme, $\alpha = 0,5$ et $K = 4,1$.

FIG. 6.2 – Convergence de Z_0 pour une grille de discrétisation non uniforme, $\alpha = 0,5$ et $K = 4,1$.

FIG. 6.3 – Convergence de Y_0 pour une grille de discrétisation non uniforme, $\alpha = 0, 1$ et $K = 4, 1$.

FIG. 6.4 – Convergence de Z_0 pour une grille de discrétisation non uniforme, $\alpha = 0, 1$ et $K = 4, 1$.

6.2.4 Comparaison des différentes grilles de discrétisation

Dans cette partie nous allons vérifier si la non uniformité de la grille de discrétisation temporelle apporte un gain numérique par rapport à la grille uniforme habituelle. Tous les exemples numériques seront donc traités avec une grille uniforme puis une grille non uniforme de paramètre $K = 0, 1$. Pour cette dernière, ρ_t est définie comme la projection sur la boule

$$B\left(0, \frac{2\sqrt{e}}{(T-t)^{\frac{2-\alpha}{4}}}\right).$$

Choix de α_M

Nous prenons $\alpha = 0,5$. Les figures 6.5 et 6.6 sont obtenues pour $\alpha_M \in \{2,5;3\}$. Il semble clair que $\alpha_M = 2,5$ n'est pas suffisant pour observer une convergence. Dans toute la suite nous fixerons donc $\alpha_M = 3$. Notons que lorsque le générateur a une croissance linéaire, $\alpha_M = 2,5$ suffit pour assurer la convergence du schéma (c.f. [63]) : le cadre quadratique est donc plus exigeant en ce qui concerne le réglage des paramètres. Les deux types de grille semblent se comporter de la même façon pour l'approximation de Y_0 tandis que la grille uniforme est légèrement meilleure en ce qui concerne Z_0 . Cela tendrait à prouver que l'introduction de cette grille non uniforme se justifie théoriquement mais pas dans la pratique.

Rôle de la projection

Nous avons vu que l'approximation de g par une fonction lipschitz implique que Z^n est borné. Il est alors envisageable de modifier la projection ρ pour en tenir compte. Ainsi, la projection 1 désignera par la suite la projection définie précédemment tandis que pour la projection 2, ρ_t est définie comme la projection sur la boule

$$B\left(0, \frac{2\sqrt{e}}{(T-t)^{\frac{2-\alpha}{4}}} \wedge N\right).$$

Les figures 6.7 et 6.8 permettent de comparer ces deux projections : la projection 2 permet d'obtenir de meilleurs résultats pour Y_0 tandis que pour Z_0 la conclusion est moins nette. Encore une fois, il n'y a pas de différence majeure entre les deux types de grille.

Rôle de l'approximation de g

L'approximation de g par une fonction N -lipschitz crée un biais dans notre approximation. Il peut être intéressant de quantifier l'erreur commise entre la solution de notre EDSR initiale et la solution de l'EDSR associée à la condition terminale g_N . Les tableaux qui suivent donnent les erreurs relatives en pourcentage sur Y_0 et Z_0 lorsque $N = n^b$ et $N = 10n^b$.

$\alpha = 0,5$

| j | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|------------------|--------|--------|--------|--------|--------|--------|--------|--------|--------|
| n^b | 1,175 | 1,175 | 1,380 | 1,454 | 1,622 | 1,747 | 1,906 | 2,052 | 2,239 |
| erreur Y_0 (%) | 8,22 | 8,22 | 4,85 | 4,09 | 2,85 | 2,24 | 1,69 | 1,33 | 1,00 |
| erreur Z_0 (%) | 3,91 | 3,91 | 2,28 | 1,92 | 1,34 | 1,05 | 0,79 | 0,62 | 0,47 |
| $10n^b$ | 11,75 | 11,75 | 13,80 | 14,54 | 16,22 | 17,47 | 19,06 | 20,52 | 22,39 |
| erreur Y_0 (%) | < 0,01 | < 0,01 | < 0,01 | < 0,01 | < 0,01 | < 0,01 | < 0,01 | < 0,01 | < 0,01 |
| erreur Z_0 (%) | < 0,01 | < 0,01 | < 0,01 | < 0,01 | < 0,01 | < 0,01 | < 0,01 | < 0,01 | < 0,01 |

$\alpha = 0,1$

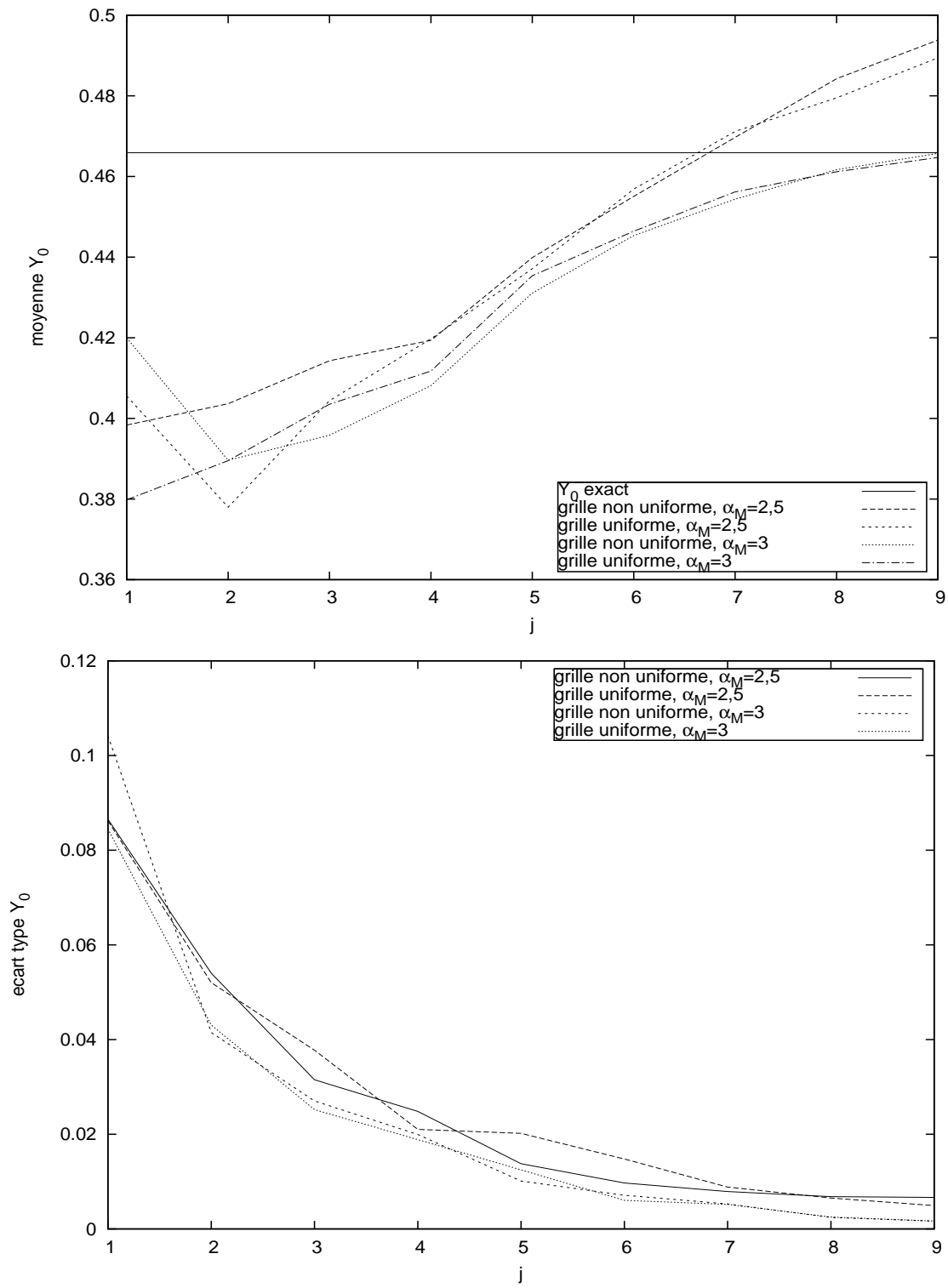
| j | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---------------------------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| n^b | 1,330 | 1,330 | 1,768 | 1,938 | 2,351 | 2,679 | 3,125 | 3,562 | 4,155 |
| erreur relative Y_0 (%) | 21,11 | 21,11 | 14,71 | 13,08 | 10,19 | 8,61 | 7,05 | 5,95 | 4,88 |
| erreur relative Z_0 (%) | 12,86 | 12,86 | 8,79 | 7,78 | 6,01 | 5,05 | 4,12 | 3,47 | 2,83 |
| $10n^b$ | 13,30 | 13,30 | 17,68 | 19,38 | 23,51 | 26,79 | 31,25 | 35,62 | 41,55 |
| erreur relative Y_0 (%) | 1,09 | 1,09 | 0,76 | 0,68 | 0,53 | 0,45 | 0,37 | 0,31 | 0,26 |
| erreur relative Z_0 (%) | 0,63 | 0,63 | 0,44 | 0,39 | 0,30 | 0,26 | 0,21 | 0,18 | 0,15 |

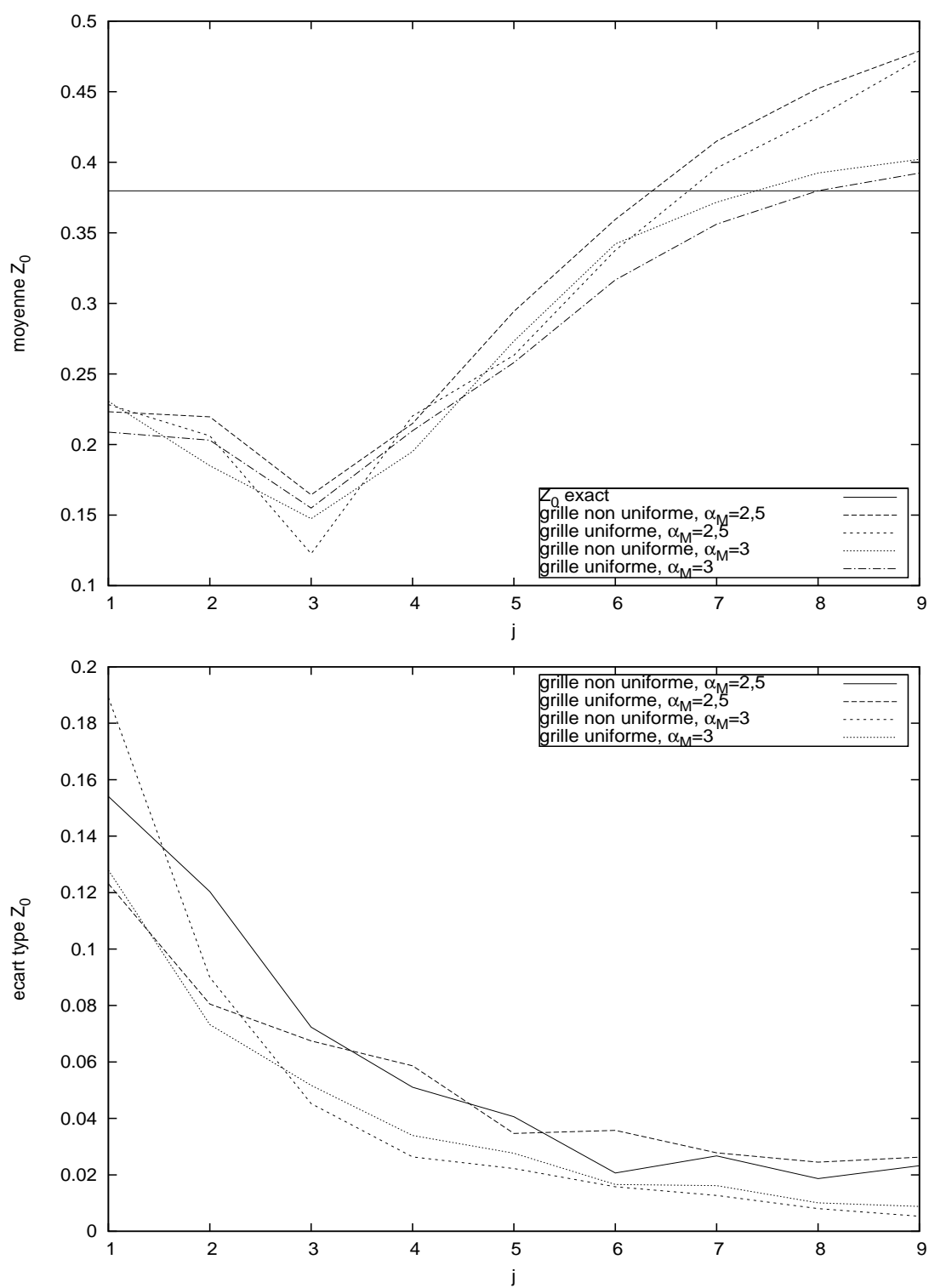
Nous pouvons observer que l'erreur relative est fortement atténuée lorsque l'on prend $N = 10n^b$. De plus, nous remarquons que lorsque $\alpha = 0,5$ et $N = n^b$ alors

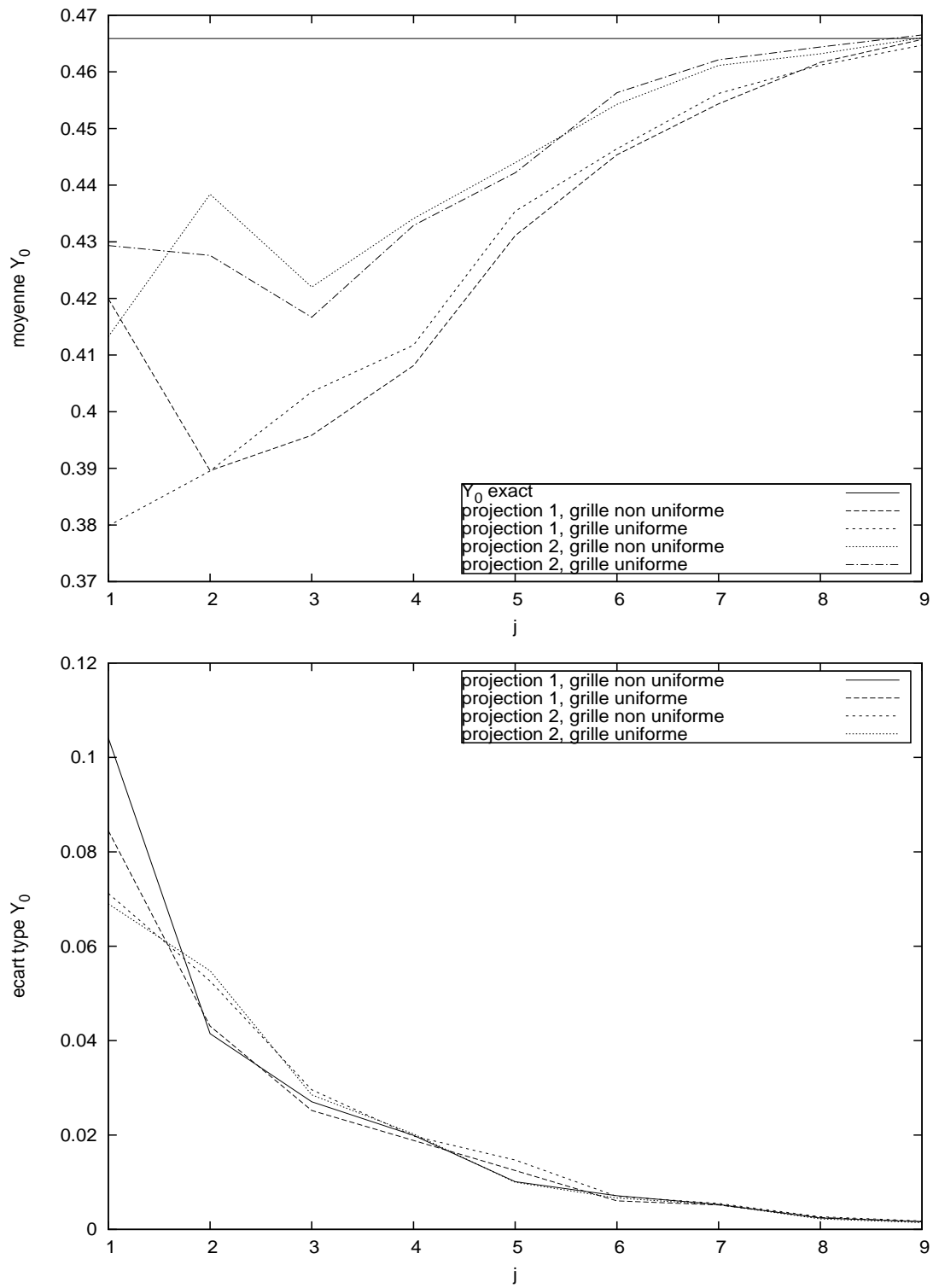
$$\frac{2\sqrt{e}}{(T-t)^{\frac{2-\alpha}{4}}} \wedge N = N$$

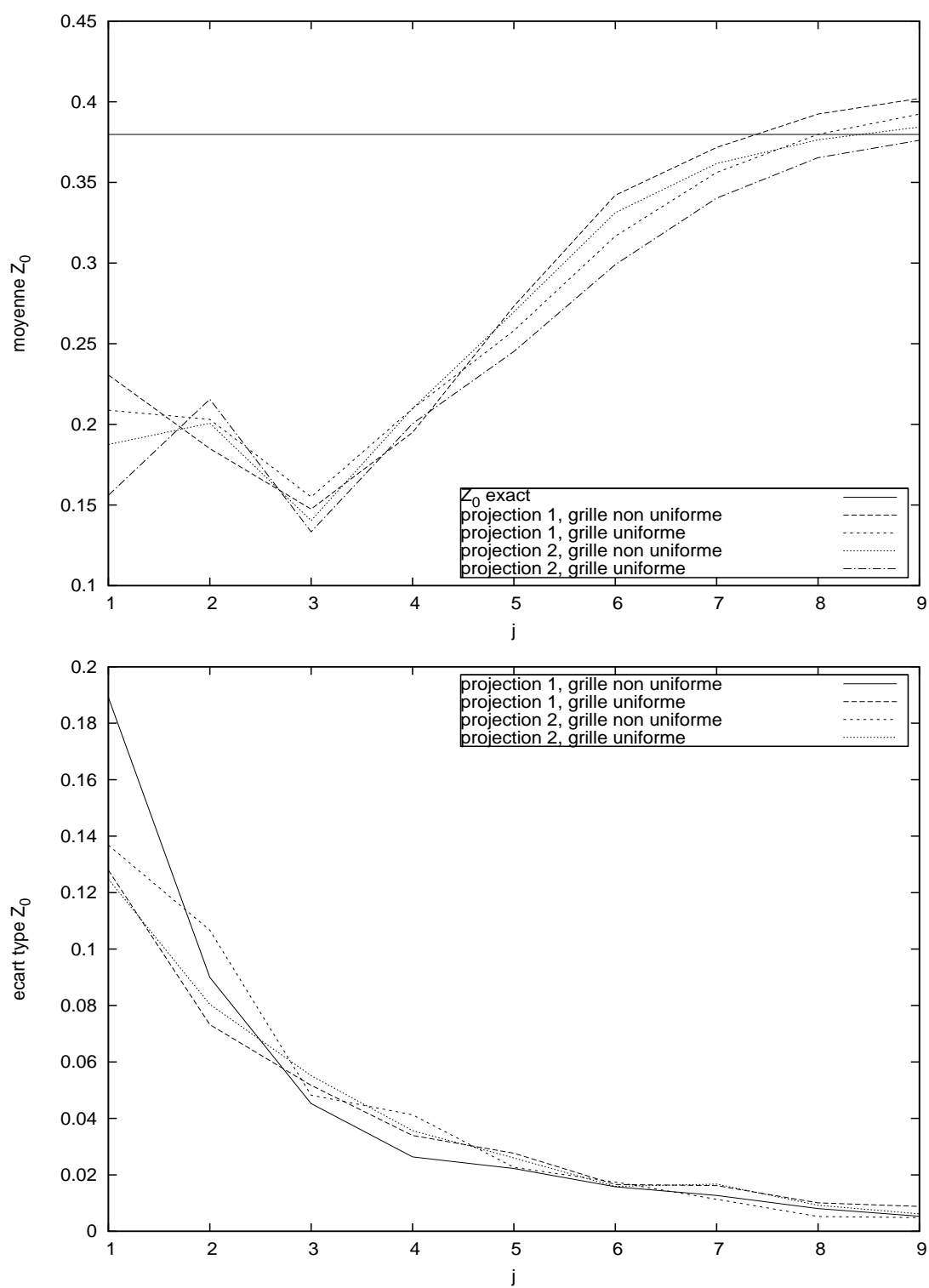
pour $j \in \{1, \dots, 9\}$. Ainsi, la projection 2 est uniforme en temps pour les figures 6.7 et 6.8 : cela pourrait expliquer que la grille non uniforme n'est pas plus efficace que la grille uniforme. Néanmoins, les figures 6.9 et 6.10, obtenues pour $\alpha = 0,5$, montrent que prendre $N = 10n^b$ est néfaste pour la convergence. De plus, il n'y a toujours pas de différence majeure entre les deux types de grille. Lorsque $\alpha = 0,1$, la figure 6.11 montre cette fois que, contrairement à ce que les résultats théoriques laissent penser, la grille uniforme est meilleure que la grille non uniforme.

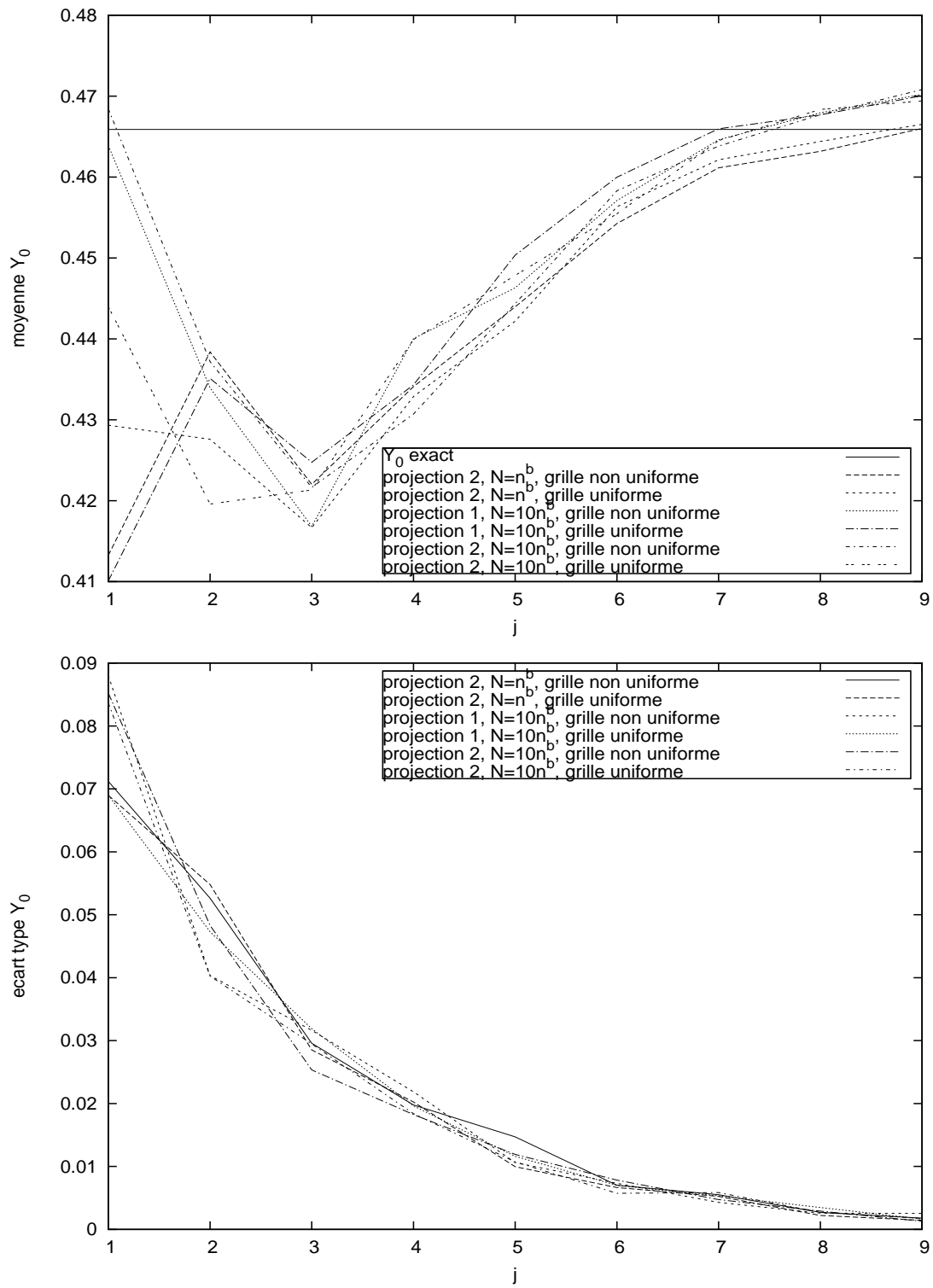
Pour terminer, nous allons étudier numériquement ce qui se passe lorsque g n'est pas approchée par une fonction lipschitz. La figure 6.12 est obtenue pour $\alpha = 0,5$, $N = n^b$ et une grille uniforme tandis que pour la figure 6.13 nous avons pris $\alpha = 0,1$, $N = 10n^b$ tout en conservant une grille uniforme. Nous pouvons observer que l'approximation ou non de g influe peu sur les résultats. Cela peut être dû au fait que g est α -höldérienne uniquement en un point.

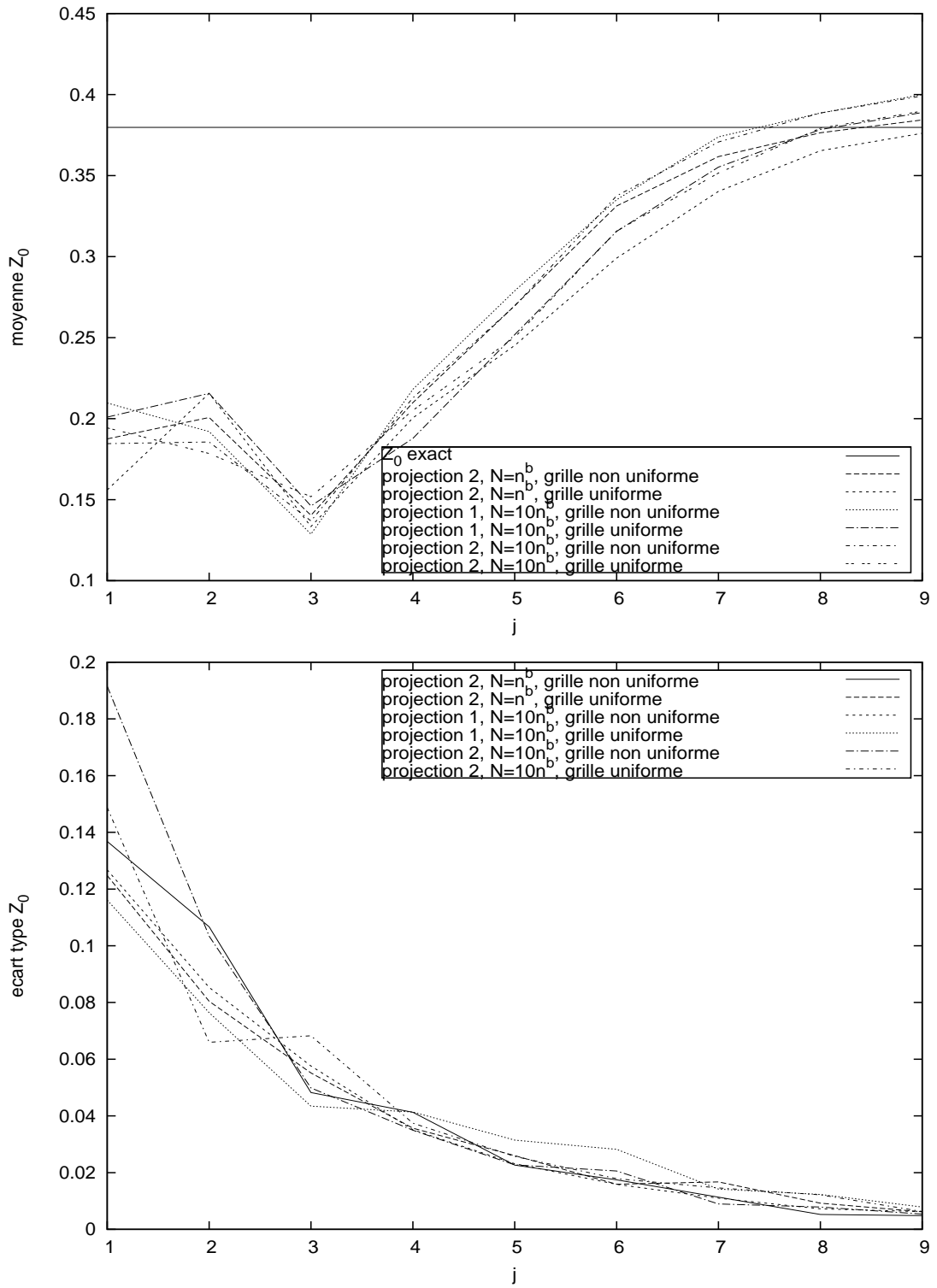
FIG. 6.5 – Convergence de Y_0 pour $\alpha_M = 2,5$ et $\alpha_M = 3$.

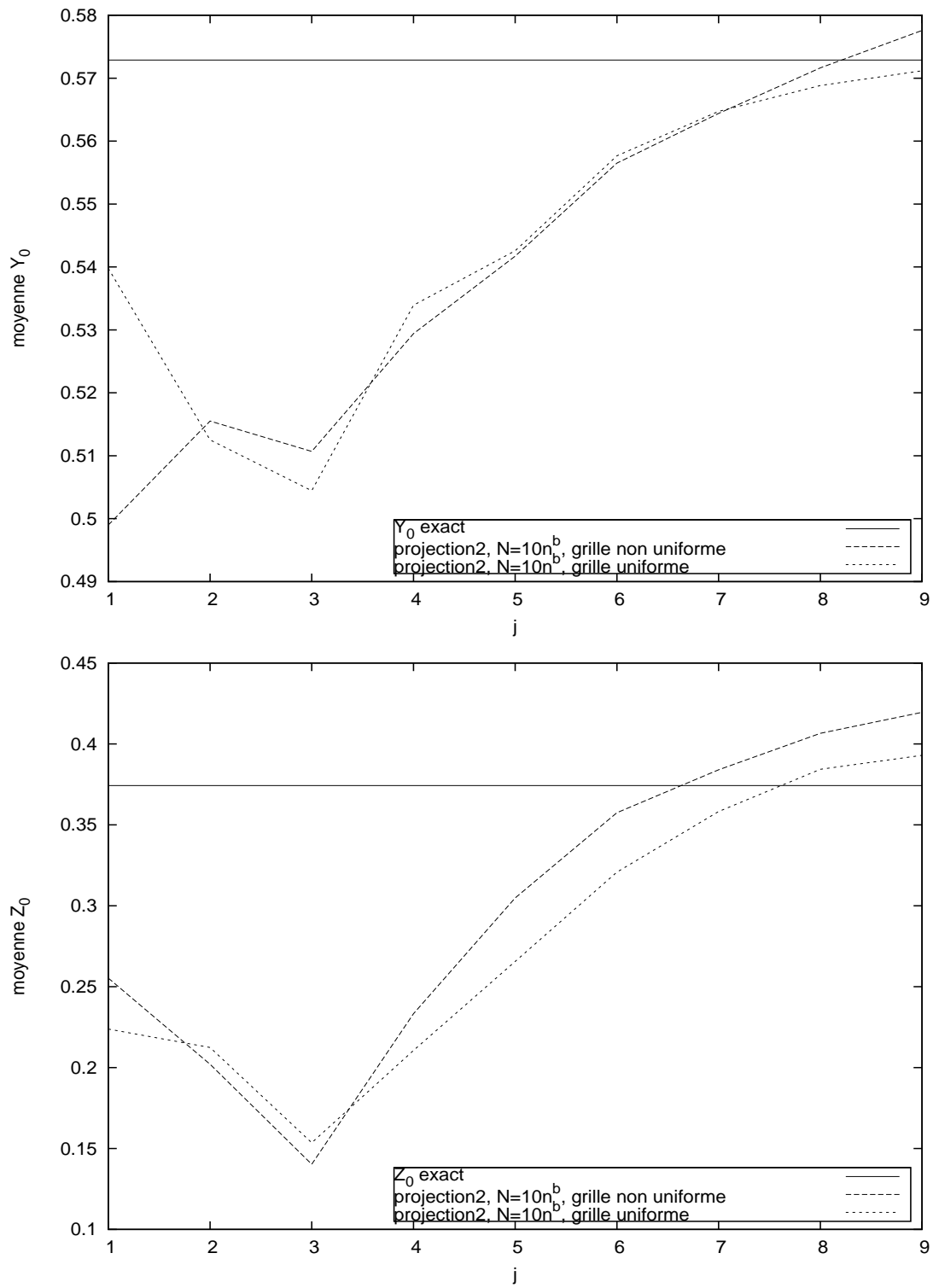
FIG. 6.6 – Convergence de Z_0 pour $\alpha_M = 2,5$ et $\alpha_M = 3$.

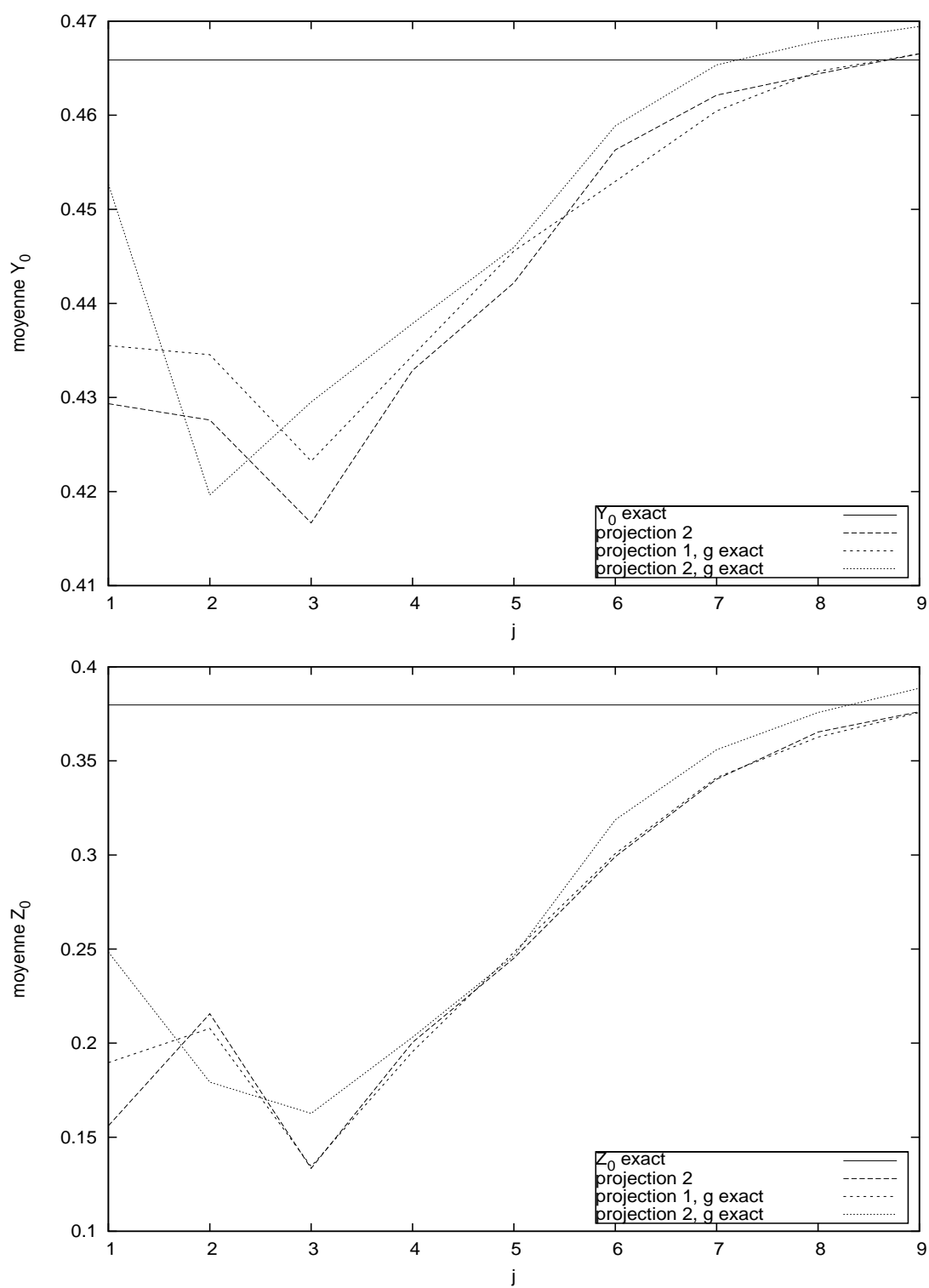
FIG. 6.7 – Convergence de Y_0 pour différentes projections.

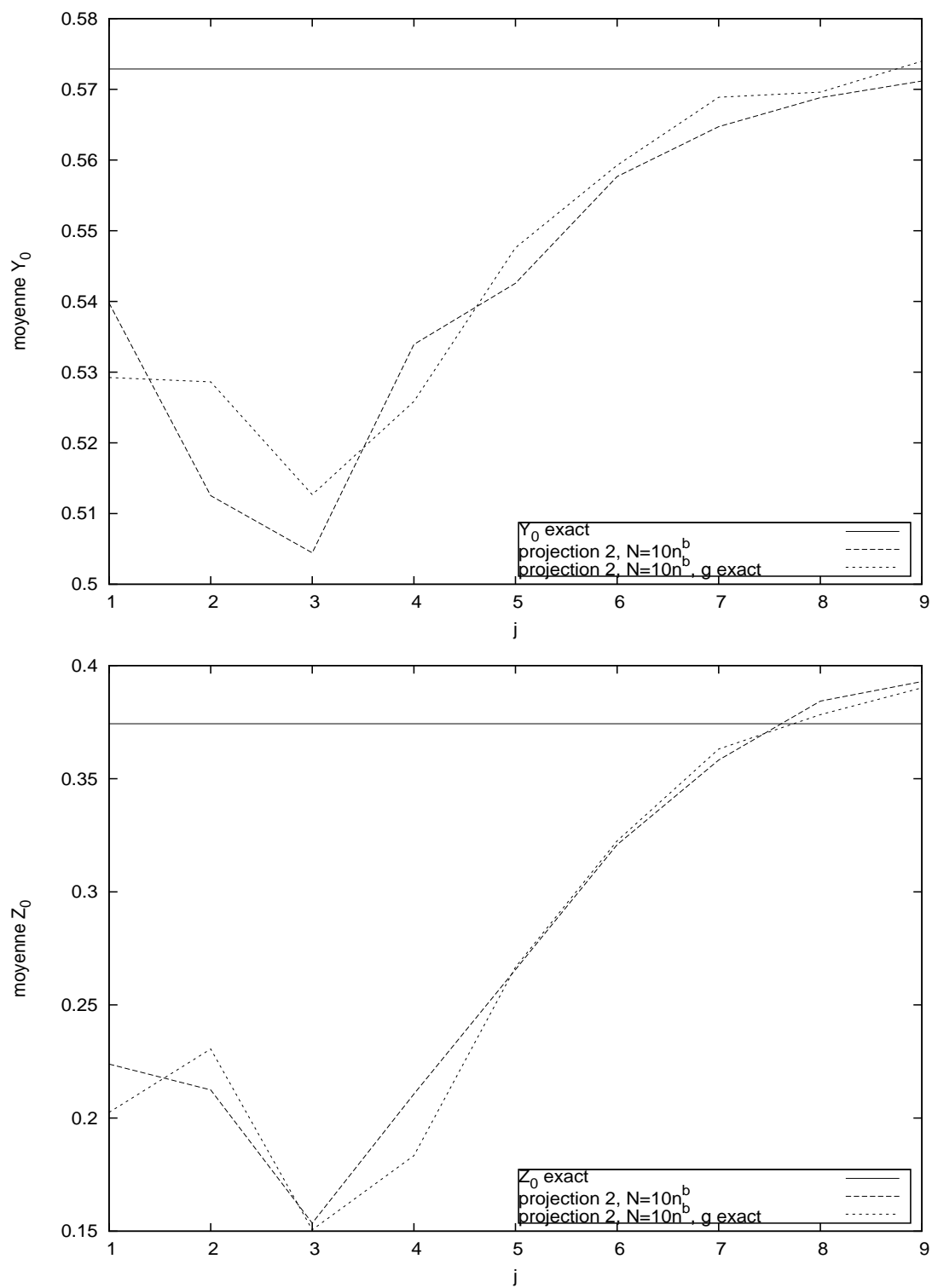
FIG. 6.8 – Convergence de Z_0 pour différentes projections.

FIG. 6.9 – Convergence de Y_0 pour différentes projections.

FIG. 6.10 – Convergence de Z_0 pour différentes projections.

FIG. 6.11 – Convergence de Y_0 et de Z_0 lorsque $\alpha = 0, 1$.

FIG. 6.12 – Convergence de Y_0 et de Z_0 lorsque la condition terminale g n'est pas approchée ($\alpha = 0,5$).

FIG. 6.13 – Convergence de Y_0 et de Z_0 lorsque la condition terminale g n'est pas approchée ($\alpha = 0, 1$).

Comparaison avec la méthode de troncature du générateur

Lorsque l'on prend pour Z une projection uniforme en temps, cela revient à considérer une EDSR dont le générateur est tronqué pour la variable z . Nous retombons alors sur la stratégie développée dans l'article d'Imkeller et dos Reis [56]. Nous avons déjà vu que la projection 2 pour $\alpha = 0,5$ était, dans les faits, une projection uniforme en temps. Dans notre cas la troncature est une puissance de n tandis que les résultats théoriques de l'article [56] nécessitent une puissance de $\log n$ afin d'obtenir la convergence du schéma. Néanmoins nous avons déjà observé avec les figures 6.7 et 6.8 que la convergence est tout de même assurée. Regardons maintenant ce qui se passe lorsque l'on prend pour troncature $N = 10n^b$. Le tableau qui suit a été obtenu pour une grille uniforme, $\alpha = 0,5$, g non approchée par g_N et une projection uniforme sur la boule $B(0, 10n^b)$.

| j | moyenne Y_0 | écart type Y_0 | moyenne Z_0 | écart type Z_0 |
|-----|---------------|------------------|---------------|------------------|
| 1 | 0,4281 | 0,08394 | 0,2004 | 0,1288 |
| 2 | 0,4206 | 0,05040 | 0,2170 | 0,08738 |
| 3 | 0,4227 | 0,02931 | 0,1391 | 0,05993 |
| 4 | 0,4371 | 0,01614 | 0,1944 | 0,03742 |
| 5 | 0,4586 | 0,03320 | 0,3022 | 0,1158 |
| 6 | 0,8412 | 1,478 | 1,070 | 2,188 |
| 7 | 15,34 | 12,14 | 18,67 | 11,74 |
| 8 | 37,70 | 14,00 | 31,27 | 6,608 |

Ainsi, le schéma numérique consistant à tronquer le générateur de l'EDSR peut grossièrement diverger lorsque la troncature augmente trop vite.

6.3 Conclusion

L'étude numérique proposée dans ce chapitre est loin d'être exhaustive. Néanmoins, il est possible d'en dégager quelques constatations empiriques :

- L'algorithme ne converge pas suffisamment rapidement lorsque K est trop grand.
- Les règles proposées par Lemor pour faire varier les paramètres du schéma s'appliquent bien dans le cadre quadratique. Néanmoins, le seuil de convergence pour α_M semble être plus élevé que pour les EDSRs dont le générateur est à croissance linéaire.
- La projection 2 donne, de manière générale, de meilleurs résultats que la projection 1.
- La grille de discrétisation temporelle non uniforme ne fournit pas de meilleurs résultats que celle qui est uniforme.
- Le schéma de discrétisation basé sur une troncature du générateur diverge lorsque cette troncature augmente trop rapidement.

Tous les tests ont été réalisés en dimension 1 ce qui, bien entendu, n'est pas suffisant. Il conviendrait donc d'étudier le comportement du schéma pour des dimensions supérieures à 1 et notamment des grandes dimensions que l'on ne peut pas traiter avec des méthodes numériques déterministes. Il serait par exemple intéressant d'appliquer ce schéma pour résoudre une EDSR dont la partie quadratique du générateur est de la forme $\frac{|z_1|^2}{2}$ avec z_1 la première composante de z . Ces EDSRs apparaissent dans l'article [54] pour traiter un problème financier d'optimisation d'utilité exponentielle et elles sont appliquées à la couverture de risques liés à la météorologie dans l'article [26]. De plus, il serait bon d'accompagner cette étude de résultats théoriques sur la convergence du schéma, du type de ceux énoncés par Lemor dans sa thèse, afin de pouvoir valider nos constatations empiriques concernant la fixation des paramètres de convergence.

Bibliographie

- [1] S. Ankirchner, P. Imkeller, and G. dos Reis. Classical and variational differentiability of BSDEs with quadratic growth. *Electron. J. Probab.*, **12**(53):1418–1453 (electronic), 2007.
- [2] M. Arisawa. Ergodic problem for the Hamilton-Jacobi-Bellman equation. I. Existence of the ergodic attractor. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, **14**(4):415–438, 1997.
- [3] M. Arisawa. Long time averaged reflection force and homogenization of oscillating Neumann boundary conditions. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, **20**(2):293–332, 2003.
- [4] M. Arisawa and P.-L. Lions. On ergodic stochastic control. *Comm. Partial Differential Equations*, **23**(11-12):2187–2217, 1998.
- [5] D. Bakry and M. Émery. Diffusions hypercontractives. In *Séminaire de probabilités, XIX, 1983/84*, volume **1123** of *Lecture Notes in Math.*, pages 177–206. Springer, Berlin, 1985.
- [6] V. Bally. Approximation scheme for solutions of BSDE. In *Backward stochastic differential equations (Paris, 1995–1996)*, volume **364** of *Pitman Res. Notes Math. Ser.*, pages 177–191. Longman, Harlow, 1997.
- [7] V. Bally and G. Pagès. A quantization algorithm for solving multi-dimensional discrete-time optimal stopping problems. *Bernoulli*, **9**(6):1003–1049, 2003.
- [8] V. Bally, G. Pagès, and J. Printems. A quantization tree method for pricing and hedging multidimensional American options. *Math. Finance*, **15**(1):119–168, 2005.
- [9] M. Bardi and I. Capuzzo-Dolcetta. *Optimal control and viscosity solutions of Hamilton-Jacobi-Bellman equations*. Systems & Control : Foundations & Applications. Birkhäuser Boston Inc., Boston, MA, 1997. With appendices by Maurizio Falcone and Pierpaolo Soravia.
- [10] G. Barles. *Solutions de viscosité des équations de Hamilton-Jacobi*, volume **17** of *Mathématiques & Applications (Berlin) [Mathematics & Applications]*. Springer-Verlag, Paris, 1994.
- [11] G. Barles and F. Da Lio. On the boundary ergodic problem for fully nonlinear equations in bounded domains with general nonlinear Neumann boundary conditions. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, **22**(5):521–541, 2005.
- [12] G. Barles, F. Da Lio, P.-L. Lions, and P. E. Souganidis. Ergodic problems and periodic homogenization for fully nonlinear equations in half-space type domains with Neumann boundary conditions. *Indiana Univ. Math. J.*, **57**(5):2355–2375, 2008.
- [13] C. Bender and R. Denk. A forward scheme for backward SDEs. *Stochastic Process. Appl.*, **117**(12):1793–1812, 2007.
- [14] A. Bensoussan and J. Frehse. Ergodic control Bellman equation with Neumann boundary conditions. In *Stochastic theory and control (Lawrence, KS, 2001)*, volume **280** of *Lecture Notes in Control and Inform. Sci.*, pages 59–71. Springer, Berlin, 2002.
- [15] J. M. Bismut. Théorie probabiliste du contrôle des diffusions. *Mem. Amer. Math. Soc.*, **4**(167):xiii+130, 1976.
- [16] B. Bouchard, I. Ekeland, and N. Touzi. On the Malliavin approach to Monte Carlo approximation of conditional expectations. *Finance Stoch.*, **8**(1):45–71, 2004.
- [17] B. Bouchard and N. Touzi. Discrete-time approximation and Monte-Carlo simulation of backward stochastic differential equations. *Stochastic Process. Appl.*, **111**(2):175–206, 2004.
- [18] B. Boufoussi and J. van Casteren. An approximation result for a nonlinear Neumann boundary value problem via BSDEs. *Stochastic Process. Appl.*, **114**(2):331–350, 2004.

- [19] P. Briand. Quelques résultats sur les équations différentielles stochastiques rétrogrades. Habilitation à diriger des recherches, 2007.
- [20] P. Briand and F. Confortola. BSDEs with stochastic Lipschitz condition and quadratic PDEs in Hilbert spaces. *Stochastic Process. Appl.*, **118**(5) :818–838, 2008.
- [21] P. Briand, B. Delyon, Y. Hu, É. Pardoux, and L. Stoica. L^p solutions of backward stochastic differential equations. *Stochastic Process. Appl.*, **108**(1) :109–129, 2003.
- [22] P. Briand, B. Delyon, and J. Mémin. Donsker-type theorem for BSDEs. *Electron. Comm. Probab.*, **6** :1–14 (electronic), 2001.
- [23] P. Briand and Y. Hu. Stability of BSDEs with random terminal time and homogenization of semilinear elliptic PDEs. *J. Funct. Anal.*, **155**(2) :455–494, 1998.
- [24] P. Briand and Y. Hu. BSDE with quadratic growth and unbounded terminal value. *Probab. Theory Related Fields*, **136**(4) :604–618, 2006.
- [25] P. Briand and Y. Hu. Quadratic BSDEs with convex generators and unbounded terminal conditions. *Probab. Theory Related Fields*, **141**(3–4) :543–567, 2008.
- [26] S. Chaumont, P. Imkeller, and M. Müller. Equilibrium trading of climate and weather risk and numerical simulation in a Markovian framework. *Stoch. Environ. Res. Risk Assess.*, **20**(3) :184–205, 2006.
- [27] P. Cheridito and M. Stadje. BSDEs and BSDEs with non-lipschitz drivers : comparison, convergence and robustness. arXiv :1002.1755v1.
- [28] D. Chevance. Numerical methods for backward stochastic differential equations. In *Numerical methods in finance*, Publ. Newton Inst., pages 232–244. Cambridge Univ. Press, Cambridge, 1997.
- [29] F. Coquet, J. Mémin, and L. Słominski. On weak convergence of filtrations. In *Séminaire de Probabilités, XXXV*, volume **1755** of *Lecture Notes in Math.*, pages 306–328. Springer, Berlin, 2001.
- [30] F. Da Lio and O. Ley. Uniqueness results for convex hamilton-jacobi equations under $p > 1$ growth conditions on data. To appear in *Appl. Math. Optim.*
- [31] F. Da Lio and O. Ley. Uniqueness results for second-order Bellman-Isaacs equations under quadratic growth assumptions and applications. *SIAM J. Control Optim.*, **45**(1) :74–106 (electronic), 2006.
- [32] A. Debussche, Y. Hu, and G. Tessitore. Ergodic BSDEs under weak dissipative assumptions. arXiv :1004.1755v1.
- [33] F. Delarue. Estimates of the solutions of a system of quasi-linear PDEs. A probabilistic scheme. In *Séminaire de Probabilités XXXVII*, volume **1832** of *Lecture Notes in Math.*, pages 290–332. Springer, Berlin, 2003.
- [34] F. Delarue and S. Menozzi. A forward-backward stochastic algorithm for quasi-linear PDEs. *Ann. Appl. Probab.*, **16**(1) :140–184, 2006.
- [35] F. Delbaen, Y. Hu, and X. Bao. Backward sdes with superquadratic growth. *Probab. Theory Related Fields*, pages 1–48, 2010.
- [36] F. Delbaen, Y. Hu, and A. Richou. On the uniqueness of solutions to quadratic BSDEs with convex generators and unbounded terminal conditions. To appear in *Ann. Inst. Henri Poincaré Probab. Stat.*, 2009.
- [37] C. M. Delfour and J. P. Zolésio. Shape analysis via oriented distance functions. *J. Funct. Anal.*, **123**(1) :129–201, 1994.
- [38] H. Djellout, A. Guillin, and L. Wu. Transportation cost-information inequalities and applications to random dynamical systems and diffusions. *Ann. Probab.*, **32**(3B) :2702–2732, 2004.
- [39] Jr. J. Douglas, J. Ma, and P. Protter. Numerical methods for forward-backward stochastic differential equations. *Ann. Appl. Probab.*, **6**(3) :940–968, 1996.
- [40] N. El Karoui and L. Mazliak, editors. *Backward stochastic differential equations*, volume **364** of *Pitman Research Notes in Mathematics Series*. Longman, Harlow, 1997. Papers from the study group held at the University of Paris VI, Paris, 1995–1996.
- [41] N. El Karoui, S. Peng, and M. C. Quenez. Backward stochastic differential equations in finance. *Math. Finance*, **7**(1) :1–71, 1997.

- [42] M. Fuhrman, Y. Hu, and G. Tessitore. On a class of stochastic optimal control problems related to BSDEs with quadratic growth. *SIAM J. Control Optim.*, **45**(4) :1279–1296 (electronic), 2006.
- [43] M. Fuhrman, Y. Hu, and G. Tessitore. Ergodic BSDEs and optimal ergodic control in Banach spaces. *SIAM J. Control Optim.*, **48**(3) :1542–1566, 2009.
- [44] M. Fuhrman and G. Tessitore. The Bismut-Elworthy formula for backward SDEs and applications to nonlinear Kolmogorov equations and control in infinite dimensional spaces. *Stoch. Stoch. Rep.*, **74**(1-2) :429–464, 2002.
- [45] A. Gégout-Petit and É. Pardoux. Équations différentielles stochastiques rétrogrades réfléchies dans un convexe. *Stochastics Stochastics Rep.*, **57**(1-2) :111–128, 1996.
- [46] E. Gobet and C. Labart. Error expansion for the discretization of backward stochastic differential equations. *Stochastic Process. Appl.*, **117**(7) :803–829, 2007.
- [47] E. Gobet and C. Labart. Solving BSDE with adaptive control variate. *SIAM J. Numer. Anal.*, **48**(1) :257–277, 2010.
- [48] E. Gobet, J. P. Lemor, and X. Warin. A regression-based Monte Carlo method to solve backward stochastic differential equations. *Ann. Appl. Probab.*, **15**(3) :2172–2202, 2005.
- [49] E. Gobet, J. P. Lemor, and X. Warin. Rate of convergence of an empirical regression method for solving generalized backward stochastic differential equations. *Bernoulli*, **12**(5) :889–916, 2006.
- [50] E. Gobet and S. Maire. Sequential control variates for functionals of Markov processes. *SIAM J. Numer. Anal.*, **43**(3) :1256–1275 (electronic), 2005.
- [51] E. Gobet and A. Makhlof. L^2 -time regularity of BSDEs with irregular terminal functions. *Stochastic Process. Appl.*, **120**(7) :1105–1132, 2010.
- [52] A. Guillin, C. Léonard, L. Wu, and N. Yao. Transportation-information inequalities for Markov processes. *Probab. Theory Related Fields*, **144**(3-4) :669–695, 2009.
- [53] R. Z. Has'minskiĭ. *Stochastic stability of differential equations*, volume 7 of *Monographs and Textbooks on Mechanics of Solids and Fluids : Mechanics and Analysis*. Sijthoff & Noordhoff, Alphen aan den Rijn, 1980. Translated from the Russian by D. Louvish.
- [54] Y. Hu, P. Imkeller, and M. Müller. Partial equilibrium and market completion. *Int. J. Theor. Appl. Finance*, **8**(4) :483–508, 2005.
- [55] Y. Hu, P. Imkeller, and M. Müller. Utility maximization in incomplete markets. *Ann. Appl. Probab.*, **15**(3) :1691–1712, 2005.
- [56] P. Imkeller and G. dos Reis. Path regularity and explicit convergence rate for BSDE with truncated quadratic growth. *Stochastic Process. Appl.*, **120**(3) :348–379, 2010.
- [57] P. Imkeller, G. dos Reis, and J. Zhang. Results on numerics for FBSDE with drivers of quadratic growth. In Alexander Chiarella, Carl ; Novikov, editor, *Contemporary Quantitative Finance*, page 440. Springer, 2010. ISBN : 978-3-642-03478-7, Essays in Honour of Eckhard Platen.
- [58] N. Kazamaki. *Continuous exponential martingales and BMO*, volume 1579 of *Lecture Notes in Mathematics*. Springer-Verlag, Berlin, 1994.
- [59] P. E. Kloeden and E. Platen. *Numerical solution of stochastic differential equations*, volume 23 of *Applications of Mathematics (New York)*. Springer-Verlag, Berlin, 1992.
- [60] M. Kobylanski. Backward stochastic differential equations and partial differential equations with quadratic growth. *Ann. Probab.*, **28**(2) :558–602, 2000.
- [61] O. A. Ladyženskaja and N. N. Ural'ceva. *Équations aux dérivées partielles de type elliptique*. Traduit par G. Roos. Monographies Universitaires de Mathématiques, No. 31. Dunod, Paris, 1968.
- [62] J.-M. Lasry and P.-L. Lions. Nonlinear elliptic equations with singular boundary conditions and stochastic control with state constraints. I. The model problem. *Math. Ann.*, **283**(4) :583–630, 1989.
- [63] J. P. Lemor. *Approximation par projections et simulation de Monte-Carlo des équations différentielles stochastiques rétrogrades*. PhD thesis, École Polytechnique, 2005.
- [64] J.-P. Lepeltier and J. San Martín. Existence for BSDE with superlinear-quadratic coefficient. *Stochastics Stochastics Rep.*, **63**(3-4) :227–240, 1998.

- [65] P.-L. Lions and A.-S. Sznitman. Stochastic differential equations with reflecting boundary conditions. *Comm. Pure Appl. Math.*, **37**(4) :511–537, 1984.
- [66] J. Ma, P. Protter, J. San Martín, and S. Torres. Numerical method for backward stochastic differential equations. *Ann. Appl. Probab.*, **12**(1) :302–316, 2002.
- [67] J. Ma, P. Protter, and J. M. Yong. Solving forward-backward stochastic differential equations explicitly—a four step scheme. *Probab. Theory Related Fields*, **98**(3) :339–359, 1994.
- [68] E. J. McShane and Jr. R. B. Warfield. On Filippov’s implicit functions lemma. *Proc. Amer. Math. Soc.*, **18** :41–47, 1967.
- [69] J.-L. Menaldi and M. Robin. Reflected diffusion processes with jumps. *Ann. Probab.*, **13**(2) :319–341, 1985.
- [70] M. A. Morlais. *Équations différentielles stochastiques rétrogrades à croissance quadratique et applications*. PhD thesis, Université de Rennes 1, 2007.
- [71] G. Pagès, H. Pham, and J. Printems. Optimal quantization methods and applications to numerical problems in finance. In *Handbook of computational and numerical methods in finance*, pages 253–297. Birkhäuser Boston, Boston, MA, 2004.
- [72] G. Pagès and J. Printems. Optimal quadratic quantization for numerics : the Gaussian case. *Monte Carlo Methods Appl.*, **9**(2) :135–165, 2003.
- [73] É. Pardoux and S. Peng. Backward stochastic differential equations and quasilinear parabolic partial differential equations. In *Stochastic partial differential equations and their applications (Charlotte, NC, 1991)*, volume **176** of *Lecture Notes in Control and Inform. Sci.*, pages 200–217. Springer, Berlin, 1992.
- [74] É. Pardoux and S. G. Peng. Adapted solution of a backward stochastic differential equation. *Systems Control Lett.*, **14**(1) :55–61, 1990.
- [75] É. Pardoux and S. Zhang. Generalized BSDEs and nonlinear Neumann boundary value problems. *Probab. Theory Related Fields*, **110**(4) :535–558, 1998.
- [76] S. Peng. Backward stochastic differential equations and applications to optimal control. *Appl. Math. Optim.*, **27**(2) :125–144, 1993.
- [77] A. Richou. Ergodic BSDEs and related PDEs with Neumann boundary conditions. *Stochastic Process. Appl.*, **119**(9) :2945–2969, 2009.
- [78] R. Rouge and N. El Karoui. Pricing via utility maximization and entropy. *Math. Finance*, **10**(2) :259–276, 2000. INFORMS Applied Probability Conference (Ulm, 1999).
- [79] M. Royer. BSDEs with a random terminal time driven by a monotone generator and their links with PDEs. *Stoch. Stoch. Rep.*, **76**(4) :281–307, 2004.
- [80] A. V. Skorokhod. *Asymptotic methods in the theory of stochastic differential equations*, volume **78** of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1989. Translated from the Russian by H. H. McFaden.
- [81] R. Tevzadze. Solvability of backward stochastic differential equations with quadratic growth. *Stochastic Process. Appl.*, **118**(3) :503–515, 2008.
- [82] S. Toldo. Stability of solutions of BSDEs with random terminal time. *ESAIM Probab. Stat.*, **10** :141–163 (electronic), 2006.
- [83] S. Toldo. Corrigendum to : “Stability of solutions of BSDEs with random terminal time” [ESAIM Probab. Stat. **10** (2006), 141–163]. *ESAIM Probab. Stat.*, **11** :381–384 (electronic), 2007.
- [84] J. Zhang. A numerical scheme for BSDEs. *Ann. Appl. Probab.*, **14**(1) :459–488, 2004.
- [85] J. Zhang. Representation of solutions to BSDEs associated with a degenerate FSDE. *Ann. Appl. Probab.*, **15**(3) :1798–1831, 2005.
- [86] J. Zhang. *Some fine properties of BSDE*. PhD thesis, Purdue University, August 2001.

Résumé

Étude théorique et numérique des équations différentielles stochastiques rétrogrades

Cette thèse est composée de trois parties indépendantes. Dans un premier temps, nous étudions une nouvelle classe d'équations différentielles stochastiques rétrogrades - notées EDSRs - qui sont reliées à des conditions de Neumann semi-linéaires relatives à des phénomènes ergodiques. La particularité de ces problèmes est que la constante ergodique apparaît dans la condition au bord. Nous étudions l'existence et l'unicité de solutions pour de telles EDSRs ergodiques ainsi que le lien avec les équations aux dérivées partielles. Nous appliquons également ces résultats à des problèmes de contrôle ergodique optimal.

Dans une deuxième partie nous généralisons des travaux de P. Briand et Y. Hu publiés en 2008. Ces derniers ont prouvé un résultat d'unicité pour les solutions d'EDSRs quadratiques de générateur convexe et de condition terminale non bornée ayant tous leurs moments exponentiels finis. Nous prouvons que ce résultat d'unicité reste vrai pour des solutions qui admettent uniquement certains moments exponentiels finis. Ces moments exponentiels sont reliés de manière naturelle à ceux présents dans le théorème d'existence. À l'aide de ce résultat d'unicité nous pouvons améliorer la formule de Feynman-Kac non linéaire prouvée par P. Briand et Y. Hu.

Enfin, nous nous intéressons à la résolution numérique d'EDSRs quadratiques markoviennes dont la condition terminale est bornée. Nous estimons dans un premier temps des bornes déterministes sur le processus Z et nous précisons le théorème de Zhang portant sur la régularité des trajectoires. Nous donnons ensuite un nouveau schéma de discrétisation en temps dont la particularité est que la grille de discrétisation est non uniforme. Enfin nous obtenons une vitesse de convergence pour ce schéma. Par ailleurs, quelques simulations numériques permettent d'étudier l'efficacité de notre nouveau schéma dans un cadre pratique.

Mots clés : équations différentielles stochastiques rétrogrades, contrôle ergodique, générateur à croissance quadratique, schéma de discrétisation temporelle, formule de Feynman-Kac non linéaire.

Abstract

Theoretical and numerical study of backward stochastic differential equations

This thesis is made of three independent parts. Firstly, we study a new class of ergodic backward stochastic differential equations - EBSDEs for short - which is linked with semi-linear Neumann type boundary value problems related to ergodic phenomena. The particularity of these problems is that the ergodic constant appears in Neumann boundary conditions. We study the existence and uniqueness of solutions to EBSDEs and the link with partial differential equations. We also apply these results to optimal ergodic control problems. In a second part, we generalise a work of P. Briand and Y. Hu published in 2008. These authors have proved the uniqueness among the solutions of quadratic BSDEs with convex generators and unbounded terminal conditions which admit every exponential moments. We prove that uniqueness holds among solutions which admit some given exponential moments. These exponential moments are natural as they are given by the existence theorem. Thanks to this uniqueness result we can strengthen the nonlinear Feynman-Kac formula proved by P. Briand and Y. Hu.

Finally, we deal with the numerical resolution of Markovian quadratic BSDEs with bounded terminal conditions. We first show some bound estimates on the process Z and we specify the Zhang's path regularity theorem. Then we give a new time discretization scheme with a non uniform time net for such BSDEs and we obtain an explicit convergence rate for this scheme. We also compute some numerical simulations to study the efficiency of our scheme in a practical situation.

Keywords : Backward stochastic differential equations, ergodic control, driver of quadratic growth, time discretization scheme, nonlinear Feynman-Kac formula.